



⑯ BUNDESREPUBLIK

DEUTSCHLAND



DEUTSCHES

PATENT- UND  
MARKENAMT

⑯ Offenlegungsschrift  
⑯ DE 198 40 771 A 1

⑯ Int. Cl. 7:  
**C 12 N 9/12**  
C 07 K 19/00  
C 12 P 19/34  
C 12 N 15/62  
C 12 N 15/63  
C 12 Q 1/68

B 11

⑯ Aktenzeichen: 198 40 771.8  
⑯ Anmeldetag: 7. 9. 1998  
⑯ Offenlegungstag: 10. 2. 2000

⑯ Innere Priorität:

198 35 653. 6 06. 08. 1998

⑯ Anmelder:

Lion Bioscience AG, 69120 Heidelberg, DE

⑯ Vertreter:

BOEHMERT & BOEHMERT, 28209 Bremen

⑯ Erfinder:

Voss, Hartmut, Dr., 69221 Dossenheim, DE;  
Moeckel, Gerd, Dr., 68723 Oftersheim, DE; Kober,  
Ingo, Dr., 69251 Gaiberg, DE; Kilger, Christian, Dr.,  
69121 Heidelberg, DE

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

⑯ Thermostabiler in vitro-Komplex mit Polymeraseaktivität

⑯ Ein erfindungsgemäßer thermostabiler in vitro-Komplex zur Template-abhängigen Elongation von Nukleinsäuren umfaßt ein thermostabiles Gleitklammerprotein, welches mit einem thermostabilen Polymeraseaktivität aufweisenden Elongationsprotein verbunden ist. Ein erfindungsgemäßer thermostabiler akzessorischer in vitro-Komplex enthält ein Gleitklammerprotein und ein Kopplungsprotein und kann mit geeigneten Polymerasen kombiniert werden. Die erfindungsgemäßen in vitro-Komplexe können in Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, insbesondere z. B. der PCR-Reaktion, eingesetzt werden. Entsprechende Reagenzienkits werden ebenfalls beschrieben.

DE 198 40 771 A 1

## Beschreibung

Die Erfindung betrifft einen thermostabilen in vitro-Komplex zur Template-abhängigen Elongation von Nukleinsäuren, einen thermostabilen prokaryontischen akzessorischen in vitro Komplex sowie dafür codierende DNA-Sequenzen und Vektoren. Die Erfindung betrifft weiterhin die Verwendung der erfindungsgemäßen Komplexe in Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wie PCR-Reaktionen oder der DNA Sequenzierung, bei denen in vitro Template-abhängige DNA Synthesen erfolgt. Schließlich betrifft die Erfindung noch Reagenzienkits zur Durchführung der erfindungsgemäßen Verfahren.

10 DNA Polymerasen gehören zu einer Gruppe von Enzymen, die einzelsträngige DNA als Template für die Synthese eines komplementären DNA Stranges verwenden. Diese Enzyme spielen eine bedeutende Rolle im Nukleinsäurestoffwechsel, einschließlich der Prozesse DNA Replikation, Reparatur und Rekombination. DNA Polymerasen wurden in allen zellulären Organismen identifiziert, von bakteriellen bis zu menschlichen Zellen, in vielen Viren sowie in Bakteriophagen (Kornberg, A. & Baker, T. A. (1991) DNA Replikation WH Freeman, New York, NY). Man fasst in der Regel die Archaeabakterien und die Eubakterien zusammen zu der Gruppe der Prokaryonten, der Organismen ohne echtem Zellkern, und stellt ihnen die Eukaryonten, die Organismen mit echtem Zellkern, gegenüber. Gemeinsam sind vielen Polymerasen aus den verschiedensten Organismen oftmals Ähnlichkeiten in der Aminosäuresequenz sowie Ähnlichkeiten in der Struktur (Wang, J., Sattar, A.K.M.A.; Wang, C.C., Karam, J.D., Konigsberg, W.H. & Steitz, T.A. (1997) Crystal Structure of pol A family replication DNA polymerase from bacteriophage RB69. *Cell* 89, 1087–1099). Organismen wie der Mensch besitzen eine Vielzahl von DNA abhängigen Polymerasen, von denen jedoch nicht alle für die DNA Replikation zuständig sind, sondern einige auch DNA Reparatur durchführen. Replikative DNA Polymerasen bestehen meist aus Proteinkomplexen mit mehreren Untereinheiten, welche die Chromosomen der zellulären Organismen und Viren replizieren. Eine generelle Eigenschaft dieser replizierenden Polymerasen ist im allgemeinen eine hohe Prozessivität, das heißt, deren Fähigkeit, Tausende von Nukleotiden zu polymerisieren ohne vom DNA Template abzudissoziieren (Kornberg, A. & Baker, T. A. (1991) DNA Replikation. WH Freeman, New York, NY).

25 Bis vor kurzem waren die einzigen gut verstandenen, hochprozessiven Replikations-Mechanismen in Zellen solche, die von zellulären Replikasen (einem Protein-Komplex mit Polymerase-Aktivität) verwendet wurden, sowie der Replikationsapparat des Bakteriophagen T4 oder T7. Der dem Mechanismus zugrunde liegende Apparat beinhaltet ein Protein mit ringähnlicher Struktur, eine "Gleitklammer", welche die DNA umschließt und die katalytische Polymeraseeinheit – das "Elongationsprotein" – an die DNA bindet (Stukenberg, P.T., Studwell-Vaughan, P.S. & O'Donnell, M. (1991) Mechanism of the  $\beta$ -clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* 266, 11328–11334; Kuriyan, J. & O'Donnell, M. (1993) Sliding clamps of DNA polymerases. *J. Mol. Biol.* 234, 915–925). Die Gleitklammer ist häufig über ein oder mehrere weitere Proteine, den "Kopplungsproteinen", an das Elongationsprotein gebunden. Die dreidimensionale Struktur von verschiedenen Gleitklammerproteinen wurde bereits bestimmt:

35 – die des eukaryontischen proliferating cell nuclear antigen (PCNA) (Krishna, T.S.R., Kong, X.-P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79, 1233–1243; Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. (1996) Structure of the C-terminal region of p21WAF1/CIP1 complexed with human PCNA. *Cell* 87, 297–306),  
 40 – die der  $\beta$ -Untereinheit der Polymerase III des Eubakteriums *Escherichia coli*, (Kong, X.-P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992) Three dimensional structure of the  $\beta$  subunit of *Escherichia coli* DNA polymerase III holoenzyme; a sliding DNA clamp. *Cell* 69, 425–437)  
 45 – und die des Bakteriophagen T4 Gen45 Proteins (Kelman, Zvi, Hurwitz, J. O'Donnell, Mike (1998) Structure, 6, 121–125).

45 Die Gesamtstruktur dieser Gleitklammern ist sehr ähnlich; die Bilder der Proteingesamtstruktur von PCNA, der  $\beta$ -Untereinheit und der gp45 Ringe sind übereinandergelegt deckungsgleich (Kelman, Z. & O'Donnell, M. (1995) Structural and functional similarities of prokaryotic and eukaryotic sliding clamps. *Nucleic Acids Res.* 23, 3613–3620). Jeder Ring hat vergleichbare Dimensionen und eine zentrale Öffnung, die groß genug ist, um Duplex-DNA, also einen DNA Doppelstrang, bestehend aus den zwei komplementären DNA Strängen, zu umschließen.

50 Die Gleitklammer kann sich *in vivo* nicht selbst um die DNA herum positionieren; sie muß durch einen Klammerlader um die DNA fixiert werden. Dieser Klammerlader ist ein Proteinkomplex, der in Prokaryonten und Eukaryonten aus einer Vielzahl von Untereinheiten besteht und der beim Eubakterium *Escherichia coli*  $\gamma$ -Komplex beziehungsweise beim Menschen Replikationsfaktor C (RF-C) genannt wird (Kelman, Z. & O'Donnell, M. (1994) DNA replication – enzymology and mechanisms. *Curr. Opin. Genet. Dev.* 4, 185–195). Der Gleitklammerlader erkennt das 3'-Ende des Einzelstrang-Duplexes (Primer-Template) und positioniert die Gleitklammer in Anwesenheit von ATP um die DNA. Die Gleitklammer, welche die DNA dann umschließt, wechselwirkt mit Polymerasen, und gewährleistet so eine schnelle und prozessive DNA Synthese.

55 Im Falle des Bakteriophagen T7 wird das gleiche Ziel, eine prozessive DNA Synthese, mittels eines strukturell anderen Proteinkomplexes erreicht. Der Phage exprimiert eine eigene katalytische Polymerase, die T7 Polymerase, das Produkt des genes 5, welche mit einem Protein aus dem Wirt *Escherichia coli*, dem Thioredoxin eine Bindung eingibt und als Replikase eine hochprozessive DNA Replikation ermöglicht (Proc Natl Acad Sci USA 1992 Oct 15; 89(20): 9774–9778 Genetic analysis of the interaction between bacteriophage T7 DNA polymerase and *Escherichia coli* thioredoxin, Hima-wan JS, Richardson CC). Auch hierbei kommt es zur Klammerbildung, jedoch weist diese Klammer nicht die gleiche Struktur auf, wie z. B. im Falle des eukaryontischen PCNA.

60 65 Oft ist es nötig – wie zum Beispiel im Falle der humanen Polymerase  $\delta$  –, daß Proteine (Kopplungsproteine) die Verbindung zwischen dem katalytisch aktiven Teil der Polymerase und dem Prozessivitätsfaktor (Gleitklammer) schaffen. Beim Menschen ist dies die kleine Untereinheit der  $\delta$ -Polymerase (Zhang, S.-J., Zeng, X.-R., Zhang, P., Toomey, N.L., Chuang, R.Y., Chang, L.-S., and Lee, M.Y.W.T. (1994). A conserved region in the amino terminus of DNA polymerase  $\delta$

is involved in proliferating cell nuclear antigen binding. *J. Biol. Chem.* 270, 7988–7992). Im Falle der T7 Polymerase jedoch bindet der Prozessivitätsfaktor die katalytische Einheit der Polymerase direkt.

DNA Polymerasen werden unter anderem durch zwei Eigenschaften charakterisiert, ihre Elongationsrate, das heißt die Anzahl der Nukleotide die sie pro Sekunde in einen wachsenden DNA Strang inkorporieren können und ihre Dissoziationskonstante. Wenn die Polymerase nach jedem Inkorporationsschritt eines der Nukleotide in die wachsende Kette wieder vom Strang abdissoziert, (d. h. ein Elongationsschritt erfolgt pro Bindungsereignis), dann hat die Prozessivität den Wert 1 und die Polymerase ist nicht prozessiv. Wenn die Polymerase für wiederholte Nukleinsäureinkorporationen mit dem Strang verbunden bleibt, dann wird der Replikationsmodus als prozessiv bezeichnet und kann einen Wert von mehreren Tausend erreichen (siehe hierzu auch: *Methods in Enzymology* Volume 262, DNA Replication, Edited by J. L. Campbell, Academic press 1995, pp. 270–280).

Für die meisten *in vitro* Anwendungen, wie PCR oder Sequenzierungsprozesse ist Prozessivität eine wünschenswerte Eigenschaft, die allerdings die bislang in diesen Reaktionen eingesetzten thermostabilen Enzyme nur in geringem Maße besitzen, wohingegen die temperatursensitive, mit Thioredoxin assoziierte T7 Polymerase eine Prozessivität von einigen tausend Nukleotiden hat. Im Vergleich – die thermostabile DNA Polymerase aus *Thermus thermophilus* oder *aquaticus* haben nur eine Prozessivität von etwa 50 Nukleotiden (*Biochim Biophys Acta* 1995 Nov 7; 1264(2): 243–248 Inactivation of the 5'-3' exonuclease of *Thermus aquaticus* DNA polymerase. Merkens LS, Bryan SK, Moses RE).

Die U.S. Patente 4,683,195, 4,800,195 und 4,683,202 beschreiben die Anwendung solcher thermostabilen DNA Polymerasen in der Polymerase Kettenreaktion (PCR). In der PCR wird unter Verwendung von Primern, Template, Nukleotiden, einer DNA Polymerase eines entsprechenden Puffers und geeigneten Reaktionsbedingungen DNA neu synthetisiert. Hierbei wird die doppelsträngige Zielsequenz zumeist thermisch aufgeschmolzen, zwei Oligonukleotide werden anhybridisiert und die Komplementärsequenz mittels der Inkorporation von Nukleotiden durch die Polymerase am Template synthetisiert. Das Extensionsprodukt jedes Primers dient als Template für den nächsten Zyklus. In dieser PCR kommt es bevorzugt zur Verwendung einer thermostabilen Polymerase, welche das zyklische, thermische Aufschmelzen der DNA Stränge übersteht. So wird häufig Taq DNA Polymerase verwendet (U.S. Patent 4,965,188). Die Prozessivität der Taq DNA Polymerase ist jedoch wie oben ausgeführt relativ gering im Vergleich zur T7 Polymerase.

DNA Polymerasen finden auch in der DNA Sequenzbestimmung Anwendung (Sanger et al., *Proc. Natl. Acad. Sci., USA* 74: 5463–5467 (1997)). Häufig wird bei der Sequenzierung nach Sanger eine T7 DNA Polymerase verwendet (Tabor, S. und Richardson, C.C. *Proc. Natl. Acad. Sci., USA* 86: 4076–4080 (1989)). Später wurde das Cycle-Sequencing-Verfahren entwickelt (Murray, V. (1989) *Nucleic Acids Res.* 17, 8889), welches keincinzelsträngiges Template erfordert und die Initiierung der Sequenzreaktion mit verhältnismäßig geringen Mengen an Template erlaubt. Die hierbei zur Verwendung kommenden Polymerasen können z. B. die oben erwähnte Taq Polymerase sein (U.S. Patent 5,075,216) oder die Polymerase von *Thermotoga neapolitana* (WO 96/10640) oder andere thermostabile Polymerasen. Neuere Verfahren koppeln die exponentielle Amplifikation und die Sequenzierung eines DNA Fragmentes in einem Schritt, so daß es möglich ist, genomische DNA direkt zu sequenzieren. Eines der Verfahren, das sogenannte DEXAS-Verfahren (*Nucleic Acids Res* 1997 May 15; 25(10): 2032–2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S, *Biol Chem* 1997 Feb; 378(2): 99–105 Direct exponential amplification and sequencing (DEXAS) of genomic DNA. Kilger C, Pääbo S und DE 196 53 439.9 sowie DE 196 53 494.1) verwendet eine Polymerase mit verminderter Diskriminierungsfähigkeit gegenüber Dideoxynukleotiden (ddNTPs) im Vergleich zu Deoxynukleotiden (dNTPs) sowie einen Reaktionspuffer, zwei Primers, die preferentiell nicht equimolar vorliegen, und die oben genannten Nukleotide, um dann in mehreren Zyklen eine komplett, sequenzspezifische DNA Leiter eines Fragmentes zu erhalten welches von den Primern umspannt ist. Eine Weiterentwicklung dieses Verfahrens besteht in der Verwendung eines Polymerasegemisches, wobei eine der beiden Polymerase zwischen ddNTPs und dNTPs diskriminiert, während die zweite eine verminderde Diskriminierungsfähigkeit aufweist (*Nucleic Acids Res* 1997 May 15; 25(10): 2032–2034 Direct DNA sequence determination from total genomic DNA. Kilger C, Pääbo S).

DNA Polymerasen finden auch Anwendung in der reversen Transkription von RNA in DNA. Hierbei dient RNA als Template und die Polymerase synthetisiert einen komplementären DNA Strang. Zur Anwendung kommt hier z. B. die thermostabile DNA Polymerase aus dem Organismus *Thermus thermophilus* (Tth) (U.S. Patent 5,322,770).

Es kann zudem erwünscht sein, daß die Polymerase eine 'proof-reading' Aktivität besitzt, also eine 3'-5' Exonukleaseaktivität aufweist. Diese Eigenschaft ist insbesondere dann wünschenswert, wenn das zu synthetisierende Produkt mit einer niedrigen Fehlerrate bei der Nukleotidinkorporation hergestellt werden soll.

Die oben genannten Enzyme, die üblicherweise in PCR-Reaktionen eingesetzt werden, gehören größtenteils nicht zu den eigentlichen Replikationsenzymen, sondern es sind zumeist Enzyme, von denen man annimmt, daß sie an der DNA Reparatur beteiligt sind, weshalb deren Prozessivität relativ gering ist.

Somit war es Aufgabe der vorliegenden Erfindung, mehrere der vorgenannten Eigenschaften von Polymerasen, insbesondere hohe Prozessivität und Thermostabilität für die Verwendung in *in vitro* Reaktionen zu vereinen.

Diese Aufgabe wurde erfundungsgemäß gelöst durch die Bereitstellung eines thermostabilen *in vitro*-Komplexes zur Template-abhängigen Elongation von Nukleinsäuren, umfassend ein thermostabiles Gleitklammerprotein, welches mit einem thermostabilen Polymeraseaktivität-aufweisenden Elongationsprotein verbunden ist. Dieser Komplex kann in *in vitro*-Reaktionen, wie z. B. in PCR-Reaktionen, eingesetzt werden und weist dabei eine hohe Prozessivität auf. Vorteil ist zudem, wenn der Komplex eine geringe Fehlerquote bei der Nukleotidinkorporation aufweist, also eine erhöhte Fidelity hat. Dieser Komplex kann somit bei der Elongation, der Amplifikation und der Sequenzierung von Nukleinsäuren eingesetzt werden. Dieser Komplex ist vorzugsweise in Standard-PCR-Reaktionen einsetzbar.

Um die Verwendbarkeit eines solchen Komplexes in Standard-PCR-Reaktionen zu ermöglichen, muß eine einfache Handhabung gewährleistet sein. So wurde zwar der Replikationsapparat des Simian virus 40 bereits *in vitro* zusammenge setzt (Wage, S. & Stillman, B. (1994) *Nature*, Vol. 369, 207–221), jedoch für Standard-PCR-Reaktionen ist dieser Replikationsapparat nicht geeignet, da er unter anderem nicht thermostabil ist. Es ist somit Gegenstand der Erfindung, einen thermostabilen *in vitro*-Komplex mit Polymeraseaktivität bereitzustellen, der in Standard-PCR-Reaktionen einsetzbar ist und die eine hohe Prozessivität aufweist.

Insbesondere ist Gegenstand der Erfindung ein thermostabiler prokaryontischer *in vitro* Komplex zur Elongation von Nukleinsäuren, der ein thermostabiles Gleitklammerprotein, welches die komplementären Nukleinsäurestränge ganz oder teilweise umschließt, und ein thermostabiles, Polymeraseaktivität aufweisendes Protein umfaßt, wobei dieses Protein oder dieser Proteinkomplex mit dem Gleitklammerprotein gekoppelt ist.

5 Im Rahmen der vorliegenden Erfindung umfaßt der Begriff Polymeraseaktivität-aufweisendes Elongationsprotein auch Polymeraseaktivität-aufweisende Proteinkomplexe oder Untereinheiten solcher Komplexe, welche die Polymeraseaktivität tragen.

Thermostabil im Sinne der vorliegenden Erfindung bedeutet, daß der akzessorische Komplex mit hoher Prozessivität Nukleotide in wachsende Nukleinsäurestränge inkorporiert sowohl bei niedrigen als auch bei hohen Temperaturen, die in 10 der PCR oder einer anderen Reaktion auftreten, wie z. B. der DNA Sequenzierung.

Die PCR besteht z. B. in der Regel aus den Schritten der Denaturierung (70°C bis 98°C), dem Annealing (40°C bis 78°C) und der DNA Strangsynthese (60°C bis 76°C). Somit muß dieser Komplex mindestens zwischen ca. 60°C und ca. 70°C, insbesondere zwischen 60°C und 76°C und besonders bevorzugterweise funktionsfähig sein zwischen 40°C und 98°C. Es dürfen während der gesamten Reaktion keine irreversiblen Denaturierungsscheinungen des Komplexes oder 15 einzelner Komponenten auftreten, welche die Elongationsreaktion unterbinden oder inhibieren.

Die Kopplung zwischen Gleitklammerprotein und Polymeraseaktivität aufweisendem Elongationsprotein kann durch kovalente, aber auch durch nicht-kovalente Bindung erfolgen. In einer bevorzugten Ausführungsform sind Gleitklammerprotein und Elongationsprotein über ein Kopplungsprotein verbunden.

In einer weiteren bevorzugten Ausführungsform der vorliegenden Erfindung stammen im erfundungsgemäßen thermostabilen *in vitro* Komplex die assoziierten Proteine aus Archaeabakterien. Es ist allerdings im Rahmen der vorliegenden Erfindung ebenfalls möglich, daß die assoziierten Proteine aus Eubakterien stammen. Gegenstand der vorliegenden Erfindung ist ebenfalls ein erfundungsgemäßer thermostabiler *in vitro* Komplex, bei dem die assoziierten Proteine zum Teil 20 aus Archaeabakterien und zum Teil aus Eubakterien stammen.

Im Sinne der vorliegenden Erfindung umfaßt der Begriff prokaryontisches Protein sowohl Proteine aus Archaeabakterien und Proteine aus Eubakterien. Es ist bekannt, daß der Replikationsapparat in Archaea dem des eukaryontischen Replikationsapparates ähnlich ist, obwohl die Genomorganisation in Eukaryonten und Archaea gänzlich verschieden ist und die zelluläre Struktur der Eubakterien dem der Archaea ähnelt. (Edgell, D.R. and Doolittle, W.F. (1997). Archaea and the origin(s) of DNA replication proteins. *Cell* 89, 995-998).

Des Weiteren ist eine bevorzugte Ausführungsform der vorliegenden Erfindung ein thermostabiler prokaryontischer *in vitro* Komplex, bei dem ein Polymeraseaktivität-aufweisender Proteinkomplex vorliegt, der aus einem Kopplungsprotein und einem oder mehreren Polymeraseaktivität-aufweisenden Elongationsprotein besteht. Bevorzugt ist des Weiteren ein erfundungsgemäßer thermostabiler prokaryontischer *in vitro* Komplex, bei dem das Gleitklammerprotein eine ringförmige Struktur aufweist, welches die komplementären Nukleinsäurestränge ganz oder teilweise umschließt.

35

### Die Gleitklammer

Die folgenden Ausführungen sollen dazu dienen, die Funktion und die möglichen Erscheinungsformen des Gleitklammerproteins besser zu verstehen.

Das Gleitklammerprotein erfüllt die Funktion, die Polymeraseaktivität an die DNA zu binden. Entweder umschließt 40 das Gleitklammerprotein selbst die DNA ganz oder teilweise oder durch Assoziation an das Polymeraseaktivität-aufweisende Protein beziehungsweise an den Polymeraseaktivität-aufweisenden Proteinkomplex oder seine Untereinheit wird eine Klammer gebildet. In jedem Fall wird durch diese Klammerbildung die Prozessivität signifikant gesteigert, mindestens um eineinhalb fache.

Das heißt, der erfundungsgemäße *in vitro* Komplex besitzt eine mindestens eineinhalb fache Prozessivität im Vergleich 45 zum Elongationsprotein alleine, beziehungsweise im Vergleich zu einem Polymeraseaktivität-aufweisenden Proteinkomplex ohne Gleitklammer oder einer Untereinheit davon.

Als Gleitklammer können beispielsweise Homologe des "Proliferating Cell Nuclear Antigen"-Proteinkomplex aus dem 50 humanen Genom, oder Homologe des ebenfalls ringförmigen "β-clamp"-Proteinkomplex aus *E. coli* dienen, welche aus thermostabilen Organismen stammen und somit thermostabil sind, oder aus nicht thermostabilen Organismen stammen, und nachträglich durch Veränderung der Aminosäuresequenz thermostabil gemacht wurden (Eijsink VG, van der Zee JR, van den Burg B, Vriend G, Venema G, FEBS Lett 1991 Apr 22; 282(1): 13-16, Improving the thermostability of the neutral protease of *Bacillus stearothermophilus* by replacing a buried asparagine by leucine, Bertus Van den Burg, Gert Vriend, Oene R. Veltman, Gerard Venema, and Vincent G. H. Eijsink Engineering an enzyme to resist boiling PNAS 1998, 95: 2056-2060). Dabei kann die Gleitklammer aus mehreren Komponenten aufgebaut sein. Die im menschlichen 55 Genom identifizierte Gleitklammer besteht aus drei PCNA-Protein Komponenten (SEQ ID NO: 11) (Homotrimeres), die im *E. coli* Genom identifizierte Gleitklammer besteht aus zwei Komponenten (SEQ ID NO: 35) (Homodimeres).

Als Gleitklammer im Sinne der vorliegenden Erfindung ist hierbei insbesondere jedes Protein zu verstehen, das die funktionelle Eigenschaft der Polymeraseprozessivitätssteigerung besitzt oder der Senkung der Fehlerrate dient. Dazu kann die Gleitklammer eine ringförmige dreidimensionale Struktur aufweisen oder durch Kopplung an ein anderes Protein 60 ein ringförmige dreidimensionale Struktur bilden, durch die sie in der Lage ist, ein- und doppelsträngige DNA ganz oder teilweise zu umschließen.

Als Gleitklammer im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das

- 65 1. zu der menschlichen (Eukaryonten) PCNA Aminosäuresequenz (SEQ ID NO: 11) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignement eine mindestens 20%ige Sequenzidentität aufweist oder das
2. zu der bakteriellen β-clamp Sequenz aus *E. coli* (Eubakteria) (SEQ ID NO: 35) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignement eine mindestens 20%ige Sequenzidentität aufweist oder das
3. zu der Aminosäuresequenz des PCNA Homologen aus *Archaeoglobus fulgidus* (Archaea) (SEQ ID NO: 12) auf

einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist.

Die erfindungsgemäße Gleitklammer kann eines oder mehrere der vorgenannten Merkmale aufweisen.

Die in Abb. 1 aufgeführten Sequenzidentitäten wurden mit dem BLAST Algorithmus nach Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J., J. Mol. Biol. 215, 403-410 (1990) ermittelt und werden weiter unten erläutert.

Als Gleitklammer im Sinne der vorliegenden Erfindung sind außerdem Proteine zu verstehen, die eine oder beide der folgenden Konsensussequenzen beinhalten und an nicht mehr als vier Positionen von einer dieser Sequenzen abweichen (Abb. 4):

Region 1

(SEQ ID NO: 39)

[G A V L I M P F W]-D-X-X-X-[G A V L I M P F W]-X-X-[G A V L I M P F W]-X-[G A V L I M P F W]-X-[G A V L I M P F W]-X-X-X-X-F-X-X-Y-X-X-D und/oder

Region 2

(SEQ ID NO: 40)

[G A V L I M P F W]-X(3)-L-A-P-[K R H D E]-[G A V L I M P F W]-E.

Die Aminosäuren werden hierbei gemäß der Standard IUPAC Einbuchstaben Nomenklatur benannt und gemäß dem Prosites Pattern Beschreibungsstandard aufgeführt. Dabei sind die folgenden Aminosäuregruppen häufig zusammengefaßt:

G, A, V, I, L, I, M, P, F oder W (Aminosäuren mit nicht polaren Seitenketten)

S, T, N, Q, Y, oder C (Aminosäure mit ungeladenen polaren Seitenketten)

K, R, H, D oder E (Aminosäure mit geladenen und polaren Seitenketten)

Außerdem bedeutet X in den Sequenzprotokollen jede beliebige Aminosäure oder Insertion oder Deletion.

Aus dem in Abb. 12 dargestellten multiplen Alignment von menschlichen PCNA Homologen wurde zudem ein Hidden Markov Modell generiert. Als Gleitklammer im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, daß mit dem so erzeugten Hidden Markov Modell einen Score mehr als 20 aufweist (Abb. 12). Die Hidden Markov Modelle und die entsprechenden Scores wurden mit dem hmms Programm (Version 1.8.4, July 1997) aus dem HMMER Paket berechnet (HMMER Protein and DNA Hidden Markov Models (Version 1.8) von Sean Eddy, Dept. of Genetics, Washington University School of Medicine, St. Louis, USA).

Aus dem in Abb. 13 dargestellten multiplen Alignment von *E. coli*  $\beta$ -clamp Homologen wurde ein Hidden Markov Modell generiert. Als Gleitklammer im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score von mehr als 25 aufweist (Abb. 13).

Die Gleitklammer kann aus mehreren Komponenten aufgebaut sein, die durch eine charakteristische Bindung fest aneinander gebunden sind, so daß ein stabiler ringförmiger Molekülkomplex gebildet wird, der nicht ohne Weiteres von der DNA dissoziieren kann. Dadurch wird eine feste aber nicht kovalente Bindung an die DNA ermöglicht, die freie Verschiebbarkeit auf derselben aber nicht behindert. Die prozessivitätssteigernden Gleitklammerproteine haben zudem charakteristische lokale Moleküleigenschaften im Bereich der Wechselwirkungsregion zur DNA, welche die freie Verschiebbarkeit erleichtern und die durch in diese Region eingelagerte Wassermoleküle unterstützt werden kann.

Eine bevorzugte Ausführungsform der vorliegenden Erfindung ist weiterhin insbesondere ein thermostabiler prokaryontischer *in vitro* Komplex, wobei das Gleitklammerprotein eines der folgenden ist: AF0335 aus *Archaeoglobus fulgidus*, MJ0247 aus *Methanococcus jannaschii*, PHLA008 aus *Pyrococcus horikoshii*, MTH1312 aus *Methanobacterium thermoautotrophicum* sowie AE000761\_7 aus *Aquifex aeolicus*.

Insbesondere sind thermostabile prokaryontische *in vitro* Komplexe vom Gegenstand dieser Anmeldung umfaßt, wobei das Gleitklammerprotein eine Aminosäuresequenz gemäß SEQ ID No. 11, 12, 13, 14, 15 und 36 (*Aquifex aeolicus*) aufweist.

Der Gleitklammerlader

Des Weiteren ist als bevorzugte Ausführungsform zu verstehen, wenn zu diesem erfindungsgemäßen Komplex zusätzlich ein Gleitklammerlader gehört, welcher die Komponenten der Gleitklammer um den ununterbrochenen DNA-Strang herum zusammenfügt, bzw. diese wieder entfernt, wenn die Reaktion beendet wird. Dieser Gleitklammerlader ist bevorzugt mit dem erfindungsgemäßen *in vitro*-Komplex assoziiert.

Im Menschen besteht der Gleitklammerlader aus fünf Untereinheiten, 4 kleinen (Gleitklammerlader 1) und einer großen Untereinheit (Gleitklammerlader 2). Als Gleitklammerlader dienen erfindungsgemäß beispielsweise ein oder mehrere prokaryontische Homologe des im Menschen identifizierten "Replication Factor C"-Proteinkomplexes (Gleitklammerlader 1: SEQ ID NO: 1, 32, 33, 34 und Gleitklammerlader 2 SEQ ID NO: 6).

Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 1, 32, 33, 34) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist.

Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung ist insbesondere auch ein solches Protein zu verstehen,

das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 6) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist.

Als Klammerlader sind beispielsweise die in **Abb. 1** aufgeführten Homologe aus Archaeabakterien geeignet (SEQ ID NO: 3, 4, 5; Homologe zu Gleitklammerlader 1 und SHQ ID NO: 8, 9, 10, Homologe zu Gleitklammerlader 2). Daher ist insbesondere bevorzugt ein thermostabiler prokaryotischer *in vitro* Komplex, wobei zusätzlich ein RFC-homologes Protein bzw. Proteinkomplex anwesend ist.

Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung ist daher beispielsweise jedes Protein zu verstehen, das die folgenden Konsensussequenzen beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht (**Abb. 6**):

10

SEQ ID NO: 41

C-N-Y-X-S-[K R H D E]-I-I-X-[G A V L I M P F W]-[G A V L I M P F W]-Q-S-R-C-X-X-F-R-F-X-P-[G A V L I M P F W].

15

Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung ist daher beispielsweise auch jedes Protein zu verstehen, das die folgenden Konsensussequenzen beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht (**Abb. 7**):

20

SEQ ID NO: 42

K-X-X-L-L-X-G-P-P-G-X-G-K-T-[S T N Q Y C]-X-[G A V L I M P F W]-X-X-[G A V L I M P F W].

25

Aus dem in **Abb. 14** dargestellten multiplen Alignment von menschlichen Gleitklammerlader 1 Homologen wurde zudem ein Hidden Markov Modell generiert. Als Gleitklammerlader 1 im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr als 25 aufweist (**Abb. 14**).

Aus dem in **Abb. 15** dargestellten multiplen Alignment von menschlichen Gleitklammerlader 2 Homologen wurde zudem ein Hidden Markov Modell generiert. Als Gleitklammerlader 2 im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr als 15 aufweist (**Abb. 15**).

30

Bevorzugt ist ebenfalls, wenn ein dem *Escherichia coli*  $\gamma$ -Komplex homologes Protein als Gleitklammerlader anwesend ist.

Außerdem bevorzugt ist ein erfundungsgemäßer thermostabiler *in vitro* Komplex zur Elongation von Nukleinsäuren, bei dem neben dem Gleitklammerlader zusätzlich ATP vorliegen kann, vorzugsweise ebenfalls assoziiert.

35

### Das Kopplungsprotein

40

Das Kopplungsprotein hat die Funktion, das Elongationsprotein und das Gleitklammerprotein zu verbinden. Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist insbesondere jedes Protein zu verstehen, welches die oben beschriebene Funktion besitzt. Daher sind als Kopplungsproteine beispielsweise die in **Abb. 1** aufgeführten Homologe zu der humanen Sequenz der koppelnden Untereinheit (DPD2\_HUMAN, [SEQ ID NO: 16]) aus Archaeabakterien geeignet (SEQ ID NO: 17, 18, 19, 20, 21).

Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 16) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 18%ige Sequenzidentität aufweist.

45

Als Kopplungsprotein im Sinne der vorliegenden Erfindung ist daher jedes Protein zu verstehen, das die folgende Konsensussequenz beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht (**Abb. 5**):

SEQ ID NO: 43

50

[FL]-[G A V L I M P F W]-X-X-[G A V L I M P F W]-X-G-X(13)-[G A V L I M P F W]-X-[YR]-[G A V L I M P F W]-X-[G A V L I M P F W]-A-G-[DN]-[G A V L I M P F W]-[G A V L I M P F W]-[DS].

55

Aus dem in **Abb. 16** dargestellten multiplen Alignment von Homologen zur menschlichen koppelnden Untereinheit wurde zudem ein Hidden Markov Modell generiert. Als koppelnde Untereinheit im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr als 10 aufweist (**Abb. 16**).

Polymeraseaktivität-aufweisendes Elongationsprotein bzw. Polymeraseaktivität aufweisender Enzymkomplex bzw. Untereinheit davon

60

Einige Elongationsproteine benötigen die Anwesenheit eines Kopplungsproteins (koppelnde Untereinheit), um überhaupt Polymeraseaktivität aufzuweisen. Es ist aber auch vorstellbar, daß das Elongationsprotein direkt an die Gleitklammer bindet. Andere Elongationsproteine benötigen die Anwesenheit eines Kopplungsproteins für die Bindung an die Gleitklammer.

65

Das Elongationsprotein weist eine 5'-3'-Polymeraseaktivität oder eine Reverse-Transkriptase-Aktivität auf. Dabei sind als Elongationsproteine beispielsweise die in **Abb. 1** aufgeführten zum menschlichen Elongationsprotein (SEQ ID NO: 22) Homologe aus Archaeabakterien geeignet (SEQ ID NO: 23, 24, 25, 26).

Bevorzugterweise bindet das Elongationsprotein an das Kopplungsprotein, welches wiederum an das Gleitklammer-

protein gekoppelt ist.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist insbesondere ein solches Protein zu verstehen, das zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 22) auf einer Länge von mindestens 200 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist daher insbesondere jedes Protein zu verstehen, das die folgende Konsensussequenz beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht (Abb. 8):

SEQ ID NO: 44

D-[G A V L I M P F W]-[G A V L I M P F W]-X-X-Y-N-X-X-F-D-X-P-Y-[G A V L I M P F W]-X-X-R-A.

5

10

Aus dem in Abb. 17 dargestellten multiplen Alignment von Homologen zum menschlichen Elongationsprotein (SEQ ID NO: 22) wurde zudem ein Hidden Markov Modell generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr 20 aufweist (Abb. 17).

15

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist insbesondere auch ein solches Protein zu verstehen, das zu der archaeabakteriellen Aminosäuresequenz (SEQ ID NO: 27) auf einer Länge von mindestens 400 Aminosäuren bei einem Sequenzalignment eine mindestens 25%ige Sequenzidentität aufweist. Beispielsweise sind geeignet die aus Archaeabakterien stammenden Proteine mit der SEQ ID NO: 28, 29, 30 oder 31.

20

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist daher insbesondere auch jedes Protein zu verstehen, das die folgende Konsensussequenz beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht (Abb. 9):

25

SEQ ID NO: 45

A-[G A V L I M P F W]-R-T-A-[G A V L I M P F W]-A-[G A V L I M P F W]-[G A V L I M P F W]-T-E-G-[G A V L I M P F W]-V-X-A-P-[G A V L I M P F W]-E-G-I-A-X-V-[K R I I D E]-I.

30

Aus dem in Abb. 18 dargestellten multiplen Alignment von Homologen zum archaebakteriellen Elongationsprotein (SEQ ID NO: 27) wurde zudem ein Hidden Markov Modell generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit insbesondere jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr als 35 aufweist (Abb. 18).

35

Das Elongationsprotein kann auch eubakteriellen Ursprungs sein.

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit ebenfalls ein solches Protein zu verstehen, das zu der eubakteriellen Aminosäuresequenz (SEQ ID NO: 37) auf einer Länge von mindestens 300 Aminosäuren bei einem Sequenzalignment eine mindestens 25%ige Sequenzidentität aufweist.

40

Als Elongationsprotein im Sinne der vorliegenden Erfindung ist daher auch insbesondere jedes Protein zu verstehen, das die folgende Konsensussequenz beinhaltet und an nicht mehr als acht Positionen von dieser Sequenz abweicht (Abb. 10):

45

SEQ ID NO: 46

[G A V L I M P F W]-P-V-G-[G A V L I M P F W]-G-R-G-S-X-[G A V L I M P F W]-G-S-[G A V L I M P F W]-V-A-X-A-[G A V L I M P F W]-X-I-T-D-[G A V L I M P F W]-D-P-[G A V L I M P F W]-X-X-X-[G A V L I M P F W]-L-F-E-R-F-L-N-P-E-R-[G A V L I M P F W]-S-M-P-D.

50

Aus dem in Abb. 19 dargestellten multiplen Alignment von Homologen zum eubakteriellen Elongationsprotein (SEQ ID NO: 37) wurde zudem ein Hidden Markov Modell generiert. Als Elongationsprotein im Sinne der vorliegenden Erfindung ist somit jedes Protein zu verstehen, das mit dem so erzeugten Hidden Markov Modell einen Score mehr 20 aufweist (Abb. 19).

55

Bisher wurden einige DNA Polymerasen als Elongationsproteine ohne Kopplungsprotein und ohne Gleitklammer für Standard PCR Reaktionen eingesetzt, so z. B. DNA Polymerase I aus *Pyrococcus furiosus* (United States Patent No. 5,545,552) oder *Pyrococcus* species (European Patent Application No: 0 547 359 A1). Diese Enzyme zeichnen sich durch die Eigenschaft aus, thermostabil zu sein und häufig eine 3'-5' exonuklease Aktivität ('proof-reading' Aktivität) zu besitzen. Erst vor kurzem wurde ein Heterodimer mit Polymeraseaktivität in *Pyrococcus furiosus* entdeckt (Uemori, T., Sato, Y., Kato, I., Doi, H., and Ishino, Y. (1997). A novel DNA polymerase in the hyperthermophilic archaeon, *pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes to Cells* 2, 499-512.)

55

Bevorzugt ist im Rahmen der vorliegenden Erfindung, daß der erfundungsgemäße prokaryontische in vitro Komplex zur Elongation von Nukleinsäuren aus Proteinen besteht, die aus Archaea stammen. Das heißt, bevorzugt ist die Verwendung eines Gleitklammerproteins aus Archaeabakterien. Deswegen ist bevorzugt, daß das Elongationsprotein beziehungsweise ein Polymeraseaktivität-aufwändiger Proteinkomplex, welcher aus Elongationsprotein und Kopplungsprotein besteht, aus Archaeabakterien stammt.

60

Beispiele solcher für den erfundungsgemäßen Komplex geeigneter Proteine sind in Abb. 1 und 2 dargestellt.

Natürlich ist es möglich, durch Deletionen oder Mutationen oder durch das Anfügen von Aminosäuren die Eigenschaften dieser Proteine zu optimieren. Diese veränderten Proteine sind ebenfalls Gegenstand der vorliegenden Erfindung, solange sie den erfundungsgemäßen in vitro Komplex zur Elongation von Nukleinsäuren bilden und die oben näher spezifizierten Funktionen erfüllen.

65

Zur Veranschaulichung des erfundungsgemäßen Komplexes dient Abb. 3, die beispielhaft den Replikationsapparat in

einer möglichen Variante darstellt, wobei die Gleitkammer über eine koppelnde Untereinheit an das Elongationsprotein bindet.

Desweiteren ist bevorzugt, daß zusätzlich zu dem erfindungsgemäßen prokaryotischen akzessorischen in vitro Komplex zwei Primer anwesend sind. Primer sind in der Regel Oligonukleotide, welche an die beiden komplementären DNA 5 Strände der Zielsequenz binden, wobei sie in gegensätzlicher Orientierung, ihre 3'-Enden zueinander gerichtet, den zu amplifizierenden Abschnitt einschließen. Sie dienen als Startpunkt der Amplifikation und stellen in der Regel ein freies 3'-OH Ende für die Polymerase zur Inkorporation eines Nukleotides zur Verfügung.

Während der Verwendung des Proteinkomplexes zur Amplifikation, Elongation und Sequenzierung liegt der erfindungsgemäße Komplex bevorzugterweise in einem geeigneten Puffer vor. Geeignete Puffer sind solche, die für PCR, Sequenzierung, Nukleinsäuremarkierung und anderen in vitro Nukleinsäure Elongationsreaktionen mittels Polymerase Anwendung finden. Geeignete Puffer werden beispielweise beschrieben in Methods in Molecular Biology Vol. 15 Humana Press Totowa, New Jersey, 1993, edited by Bruce A. White.

Während der Verwendung des Proteinkomplexes zur Elongation, Amplifikation, Reversen Transkription oder/und Sequenzierung liegt zusätzlich zu dem erfindungsgemäßen Komplex ein Gemisch aus Nukleotiden vor. Deoxynukleotide 15 können aus dGTP, dATP, dTTP und dCTP ausgewählt werden, sind jedoch nicht auf diese beschränkt. Zusätzlich können auch Derivate von Deoxynukleotiden gemäß der Erfindung verwendet werden, welche als solche Deoxynukleotide definiert sind, die in der Lage sind, durch eine thermostabile DNA Polymerase in wachsende DNA Moleküle inkorporiert zu werden, die in der Reaktion synthetisiert werden. Solche Derivate schließen Thionukleotide, 7-deaza-2'-dGTP, 7-deaza-20 2'-dATP sowie Deoxyinosine Triphosphat, das auch als Ersatz Deoxynukleotid für dATP, dGTP, dTTP oder dCTP verwendet werden kann, ein, sind aber nicht auf diese beschränkt. Ebenso können markierte Deoxynukleotide verwendet werden. Alle bekannten oder/und zu dem erfindungsgemäßen Zweck geeigneten Markierungen können dabei vorliegen.

Dideoxynukleotide können aus ddGTP, ddATP, ddTTP und ddCTP ausgewählt werden, sind jedoch nicht auf diese beschränkt. Zusätzlich können auch Derivate von Dideoxynukleotiden gemäß der Erfindung verwendet werden, die als solche Dideoxynukleotide definiert sind, die in der Lage sind, durch eine thermostabile DNA Polymerase in wachsende DNA Moleküle inkorporiert zu werden, die in der Reaktion synthetisiert werden. Solche Derivate können radioaktive Dideoxynukleotide (ddATP, ddGTP, ddTTP und ddCTP) oder Dideoxynukleotide (ddATP, ddGTP, ddTTP und ddCTP), welche mit z. B. FITC, Cy5, Cy5.5, Cy7 und Texas-Rot oder anderen markiert sind, einschließen, sind aber nicht auf diese beschränkt. Im Rahmen einer erfindungsgemäßen Sequenzierung können auch markierte Deoxynukleotide zusammen mit unmarkierten Dideoxynukleotiden eingesetzt werden.

30 Ribonukleotide können aus GTP, ATP, TTP und CTP ausgewählt werden, sind jedoch nicht auf diese beschränkt. Zusätzlich können auch Derivate von Ribonukleotiden gemäß der Erfindung verwendet werden, die als solche Ribonukleotide definiert sind, die in der Lage sind, durch eine thermostabile DNA Polymerase in wachsende DNA Moleküle inkorporiert zu werden, die in der Reaktion synthetisiert werden. Solche Derivate können radioaktive Ribonukleotide (ATP, GTP, TTP und CTP) oder Ribonukleotide (ATP, GTP, TTP und CTP), welche mit z. B. FITC, Cy5, Cy5.5, Cy7 und Texas-Rot oder anderen markiert sind, einschließen, sind aber nicht auf diese beschränkt.

35 Während der Verwendung des Proteinkomplexes zur Amplifikation, Elongation und Sequenzierung kann es sich als vorteilhaft erweisen wenn bei der Reaktion eine Pyrophosphatase anwesend ist.

Ein weiterer Gegenstand der vorliegenden Erfindung ist ein thermostabiler, akzessorischer in vitro-Komplex, welcher ein Gleitklammerprotein und ein Kopplungsprotein umfaßt, wobei beide Proteine wie oben definiert sind.

40 Der erfindungsgemäße akzessorische in vitro-Komplex ist gewissermaßen als Vorstufe zu dem oben beschriebenen erfindungsgemäßen thermostabilen in vitro-Komplex zur Template-abhängigen Elongation von Nukleinsäuren zu verstehen.

Der akzessorische Komplex muß lediglich mit einem Elongationsprotein kombiniert werden, um diese in eine deutlich erhöhte Prozessivität zu verleihen. Der akzessorische in vitro-Komplex kann daher auch zusammen mit bekannten thermostabilen Polymerasen eingesetzt werden, wobei ebenfalls Nachteile dieser bekannten Proteine hinsichtlich insbesondere der Prozessivität verringert werden können.

Identifizierung der Gene, Klonierung der Gene, Expression dieser und Reinigung der Proteine der erfindungsgemäßen in vitro-Komplexe:

In der Regel können die Komplexe, bestehend aus rekombinanten Proteinen, mit den folgenden Schritten bereitgestellt 50 werden: Bereitstellung des Nukleinsäurefragments welches für das gewünschte Protein kodiert, Ligation in einen Expressionsvektor, Transformation in einen Wirt, Expression und Aufreinigung des Proteins. In Übereinstimmung mit der vorliegenden Erfindung kann es vorkommen, daß Gene, insbesondere aus den Archaeabakterien, Inteine (Proc Natl Acad Sci USA 1992 Jun 15; 89(12): 5577-5581, Intervening sequences in an Archaea DNA polymerase gene, Perler FB, Comb DG, Jack WE, Moran LS, Qiang B, Kucera RB, Benner J, Slatko BE, Nwankwo DO, Hempstead SK, et al) enthalten können, welche zunächst entfernt werden können.

Die Identifizierung weiter für den erfindungsgemäßen Komplex geeigneten Proteine kann z. B. erfolgen durch Homologiesuchen in Datenbanken, welche Genome aus Prokaryonten umfassen. Einige Programme sind hierfür geeignet, beispielsweise das Programm BLASTP und FASTA (Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nuclic Acids Res. 25: 3389-3402. W.R. Pearson & D.J. Lipman PNAS (1988) 85: 2444-2448).

Die Identifizierung kann auch dadurch geschehen, daß DNA Sonden verwendet werden, um in z. B. gesamtgenomischen Banken aus Prokaryonten nach den entsprechenden Genen zu screenen. Die hierfür nötigen experimentellen Verfahren finden sich in Maniatis et al. (Molecular Cloning: A laboratory Manual, Cold Spring Harbor Laboratory Press, 65 N.Y.(1989)).

Die Bereitstellung der gereinigten Nukleinsäure der Gene der erfindungsgemäßen Komplexe kann z. B. über deren Isolierung aus einer genomischen Bank des relevanten Organismus geschehen oder durch synthetische DNA-Herstellung, jeweils gewünschtenfalls kombiniert mit einer Amplifikation mittels PCR, unter Zuhilfenahme von Primern, wel-

che für den gewünschten Genabschnitt spezifisch sind. Übliche Verfahren sind in Maniatis et al. (Molecular Cloning: A laboratory Manual, Cold Spring Harbor Laboratory Press, N.Y. (1989)) beschrieben.

Die Gene der Proteine der erfundungsgemäßen *in vitro*-Komplexe können in einer Vielzahl von Verfahren kloniert werden und somit mittels eines Expressionsvektors zur Proteinexpression in einem Wirtsorganismus zur Verfügung gestellt werden. Gängige Verfahren sind in Maniatis et al. (Molecular Cloning: A laboratory Manual, Cold Spring Harbor Laboratory Press, N.Y. (1989)) beschrieben. Die Gene der Komplexe können hierbei z. B. zunächst in einen 'high-copy' Vektor, z. B. pUC18, pBst oder pBR322 kloniert werden und dann erst in einen prokaryontischen Expressionsvektor, z. B. pTrc99 umkloniert werden, oder aber direkt in einen prokaryontischen Expressionsvektor kloniert werden. Unter Vektoren sind hierbei Nukleinsäuren zu verstehen, die in der Lage sind, ein anderes Nukleinsäuremolekül in oder zwischen verschiedenen Organismen, bzw. genetischen Hintergründen zu transportieren. Sie haben in der Regel die Fähigkeit der autonomen Replikation und/oder der Expression (Expressionsvektoren) des operativ verbundenen Nukleinsäuremoleküls. "Operativ verbunden" bedeutet, daß das transportierte Nukleinsäuremolekül so mit dem Vektor verbunden ist, daß es unter der Transkriptions- und Translationskontrolle von Expressionskontrollsequenzen des Vektors steht und in einer Wirtszelle exprimiert werden kann. Bakterielle Expressionssysteme, die bevorzugte Verwendung derer, sowie eine Auswahl an Vektorsystemen ist z. B. in 'Gene Expression Technology', (Meth. Enzymol., Vol 185, Goeddel, Ed., Academic Press, N.Y. (1990)) beschrieben. Geeignete Vektoren für die vorliegende Erfahrung sollten unterschiedlich starke Expression der Proteine dadurch ermöglichen, daß sie einige oder alle der folgenden Eigenschaften besitzen: (1) Promotoren, oder Transkriptionsinitiationsstellen, entweder unmittelbar neben dem Start des Proteins oder als Fusionsprotein, (2) Operatoren die verwendet werden können Genexpression an oder aus zu schalten, (3) ribosomale Bindungsstellen für eine verbesserte Translation, und (4) Terminationsstellen für die Transkription oder Translation, die zu verbesselter Stabilität führen.

Expressionsvektoren, die mit eukaryontischen Zellen, bevorzugterweise mit Vertebratenzellen kompatibel sind, können auch Verwendung finden. Einige bekannte Vektoren sind pSVL und pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), und pTDT1 (ATCC 31255). Die Verwendung eines retrovirusalen Expressionsvektors ist auch möglich.

Weitere Gegenstände der vorliegenden Erfahrung sind daher DNA-Sequenzen, welche für die erfundungsgemäßen thermostabilen *in vitro*-Komplexe bzw. akzessorischen *in vitro*-Komplexe codieren, sowie entsprechende Vektoren, vorzugsweise Expressionsvektoren.

Die erfundungsgemäßen Vektoren enthalten mindestens das Gen für das Glicitkammcrprotein und vorzugsweise mindestens ein Gen für ein Kopplungs- oder/und Elongationsprotein, wie sie jeweils oben definiert sind.

Es ist dabei im Rahmen der Erfahrung bevorzugt, wenn der Vektor neben den darin bereits enthaltenen DNA-Sequenzen noch geeignete Restriktionsschnittstellen und ggf. Polylinker zur Insertion weiterer DNA-Sequenzen enthält. Besonders bevorzugt ist es, wenn die räumliche Anordnung von bereits enthaltener DNA-Sequenz und zusätzlicher Insertionsstelle nach Expression zur Ausbildung eines Fusionsproteins führt.

Ebenfalls bevorzugt ist es, wenn der erfundungsgemäße Vektor Promotor- oder/und Operatorbereiche enthält, wobei es besonders bevorzugt ist, wenn derartige Promotor- oder/und Operatorbereiche induzierbar oder reprimierbar sind. Dadurch wird eine Steuerung der Expression in Wirtszellen erheblich vereinfacht und kann besonders effizient gestaltet werden.

Derartige Promotor/Operatorbereiche können in einem Expressionsvektor auch mehrfach vorkommen, so daß eine ggf. unabhängige Expression mehrerer DNA-Sequenzen unter Verwendung nur eines Expressionsvektors ermöglicht wird.

Ein besonders bevorzugter Vektor enthält im Rahmen der Erfahrung DNA-Sequenzen codierende für alle Bestandteile eines erfundungsgemäßen *in vitro*-Komplexes oder akzessorischen *in vitro*-Komplexes.

Ein weiterer Gegenstand der vorliegenden Erfahrung ist eine Wirtszelle, enthaltend einen oder mehrere erfundungsgemäße(n) Vektor(en), wobei in dieser Wirtszelle unter geeigneten Bedingungen die Expression zu Proteinen erfolgen kann. Geeignete Bedingungen schließen beispielsweise Anwesenheit eines Induktors oder eines Derepressors ein.

Für die Transformation, Phageninfektion und Zellkultur existieren Standardprotokolle in Maniatis et al. (Molecular Cloning: A laboratory Manual, Cold Spring Harbor Laboratory Press, N.Y.). Aus der Vielzahl der vorhandenen *E. coli* Stämme, die zur Transformation geeignet sind, sind die bevorzugten JM101 (ATCC No. 33876), XL1 (Stratagene), RRI (ATCC No. 31343) und BL21 (Pharmacia). Proteinexpression kann z. B. auch geschehen unter Zuhilfenahme des *E. coli* Stammes INVαF' (Invitrogen).

Die Transformanten werden gemäß den geeigneten Wachstumsbedingungen des Wirtsstammes kultiviert. So werden die meisten *E. coli* Stämme z. B. in LB Medium kultiviert, bei 30°C bis 42°C bis zur logarithmischen oder stationären Wachstumsphase. Die Proteine können aus einer transformierten Kultur gereinigt werden, wobei dies entweder aus einem Zellpellet, nach Zentrifugation, oder aber aus der Kulturlösung geschehen kann. Sofern die Proteine aus dem Zellpellet gereinigt werden, werden die Zellen in einem geeigneten Puffer resuspendiert und mittels Sonifizierung, enzymatischer Behandlung oder Einfrieren und Auftauen aufgebrochen. Sofern die Reinigung aus der Kultursuspension geschieht, das heißt alleine oder über ein Fusionsprotein, wird der Überstand von den Zellen mittels bekannter Verfahren, wie Zentrifugation abgetrennt.

Die Separation und Reinigung der Proteine der erfundungsgemäßen Komplexe entweder aus dem Überstand der Kulturlösung oder aus dem Zellextrakt kann durch bekannte Separations- oder Reinigungsverfahren geschehen. Dicse Methoden sind z. B. solche die Löslichkeiten betreffen, wie Salzfällungen, und Lösungsmittelfällungen, Methoden die sich die unterschiedlichen Molekulargewichte zu Nutze machen, wie Dialyse, Ultrafiltration, Gelfiltration, und SDS-Polyacrylamideelektrophorese, Methoden die sich die unterschiedlichen Ladungen zu Nutze machen, Ionenauftauschchromatographie, Methoden die sich die unterschiedlichen Hydrophobizitäten zu Nutze machen, wie reverse-phase HPLC (High Performance Liquid Chromatography), Methoden die sich bestimmte Affinitäten zu Nutze machen, wie Affinitätschromatographie, und Methoden die sich Unterschiede im Isoelektrischen Punkt zu Nutze machen, wie isoelektrische Fokussierung. Es ist auch vorstellbar, daß Zellextrakte, entweder aus dem Organismus, welcher das Gen des akzessori-

schen Komplexes trägt, gemacht werden können, welcher die erfindungsgemäße Aufgabe erfüllt, oder aus dem rekombinanten Wirtsorganismus, z. B. E. coli. Mit diesen Extraktionen könnte man unter Umständen andere Reinigungs-schritte umgehen.

Die oben beschriebenen Verfahren können in einer Vielzahl von Kombinationen eingesetzt werden um die Proteine des *in vitro*-Komplexes bereitzustellen.

Der erfindungsgemäße thermostabile prokaryotische akzessorische *in vitro* Komplexes kann zur Elongation von Nukleinsäuren verwendet werden, z. B. zur Polymerase-Ketten-Reaktion, DNA Sequenzierung, zur Markierung von Nukleinsäuren und anderen Reaktionen, die die Nukleinsäure *in vitro* Synthese beinhalten. Auch möglich ist die Anwendung in der reversen Transkription, wobei entweder der erfindungsgemäße Komplex selbst reverse Transkriptase Aktivität besitzt, oder aber zusätzlich ein geeignetes Enzym hinzugefügt wird, das reverse Transkriptase Aktivität besitzt.

Ein weiterer Gegenstand der vorliegenden Erfindung ist daher ein Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wobei die Nukleinsäure nötigenfalls denaturiert, mit mindestens einem Primer unter Hybridisierungsbedingungen versehen wird, wobei der Primer komplementär zu einem flankierenden Bereich einer gewünschten Nukleinsäuresequenz des Templatestrangs ist und mit Hilfe einer Polymerase in Anwesenheit von Nukleotiden eine Primerelongation erfolgt, wobei als Polymerase ein erfindungsgemäßer thermostabiler *in vitro*-Komplex eingesetzt wird.

Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, bei denen die Elongation ausgehend von einem Primer erfolgt, der an die Template-Nukleinsäure anhybridisiert wurde und ein freies 3'-OH-Ende für die Elongation zur Verfügung stellt, sind dem Fachmann bekannt. Zur Amplifikation wird insbesondere eine Polymerasekettenreaktion durchgeführt. Hierbei wird normalerweise von einer doppelsträngigen DNA-Sequenz ausgegangen, von welcher ein bestimmter Zielbereich amplifiziert werden soll. Hierbei werden zwei Primer eingesetzt, welche zu die Zielsequenz flankierenden Bereichen auf jeweils einem Teilstrang des DNA-Doppelstranges komplementär sind. Zur Anhybridisierung der Primer werden allerdings die DNA-Doppelstränge zuerst denaturiert, insbesondere thermisch aufgeschmolzen. Nach Anhybridisierung der Primer erfolgt eine Elongation mittels der Polymerase, daraufhin wird nochmals denaturiert und damit die neu gebildeten DNA-Stränge von den Template-Strängen getrennt, woraufhin für einen weiteren Elongationszyklus neben den ursprünglichen Templatesträngen auch die im ersten Schritt gebildeten Nukleinsäurestränge als Template zur Verfügung stehen, diese jeweils erneut mit Primern hybridisiert werden und eine erneute Elongation stattfindet. Diese Vorgehensweise wird zyklisch durchgeführt unter jeweils thermischer Denaturierung als Zwischenschritte.

Zur erfindungsgemäß bevorzugten Reversen Transkription von RNA in DNA wird ebenfalls ein erfindungsgemäßer thermostabiler *in vitro*-Komplex eingesetzt, wobei dessen Elongationsprotein eine Reverse Transkriptase-Aktivität aufweist. Diese Reverse Transkriptase-Aktivität kann die einzige Polymeraseaktivität des Elongationsproteins sein, kann aber auch zusätzlich zu einer vorhandenen 5'-3'-DNA-Polymeraseaktivität vorliegen.

Ein weiteres bevorzugtes erfindungsgemäßes Verfahren betrifft die Sequenzierung von Nukleinsäuren ausgehend von einem Primer, der zu einem der zu sequenziierenden Nukleinsäure benachbarten Bereich komplementär ist, wobei wiederum eine Template-abhängige Elongation oder aber bei Sequenzierung einer RNA eine Reverse Transkription unter Verwendung von Deoxynukleotiden und Dideoxynukleotiden gemäß der Methode von Sanger durchgeführt wird. Als Deoxynukleotide oder Dideoxynukleotide werden im Rahmen dieser bevorzugten Ausführungsform auch die oben beschriebenen jeweiligen Derivate als geeignet angesehen. Insbesondere ist es für die erfindungsgemäßen Verfahren zur Elongation von Nukleinsäuren bevorzugt, daß die gebildeten Nukleinsäuren markiert werden. Hierzu ist es möglich, markierte Primer oder/und markierte Deoxynukleotide oder/und markierte Dideoxynukleotide oder/und markierte Ribonukleotide oder jeweils entsprechende Derivate, wie sie oben bereits beispielweise beschrieben sind, einzusetzen.

Wiederum ein weiterer Gegenstand der vorliegenden Erfindung ist ein Verfahren zur Markierung von Nukleinsäuren durch Einfügung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, wobei als Polymerase ein erfindungsgemäßer thermostabiler *in vitro*-Komplex eingesetzt wird.

Ein solches Verfahren, allgemein Nick-Translation genannt, ermöglicht eine einfache Markierung von Nukleinsäuren. Alle oben bereits beschriebenen markierten Ribonukleotide oder Deoxyribonukleotide oder Derivate davon sind hierfür geeignet, solange die Polymerase sie als Substrat akzeptiert.

Diese erfindungsgemäße Markierung kann z. B. auch im Rahmen oder anschließend an eine PCR-Reaktion erfolgen, wofür die erfindungsgemäße Verwendung eines thermostabilen *in vitro*-Komplexes besonders günstig ist.

Gegenstand der vorliegenden Erfindung ist ebenfalls ein Kit zur Elongation oder/und Amplifikation oder/und Reversen Transkription oder/und Sequenzierung von Nukleinsäuren, wobei dieser Kit in einem oder in mehreren Behältern vorliegen kann, enthaltend

- 55 a) einen erfindungsgemäßen thermostabilen *in vitro* Komplex oder
- b) einen thermostabilen, akzessorischen *in vitro*-Komplex und ggf. separat davon ein Polymeraseaktivität aufweisendes Elongationsprotein sowie ggf. Primer, Puffersubstanzen, Nukleotide, ATP, andere Kofaktoren oder/und Pyrophosphat.

Insbesondere ist Gegenstand der vorliegenden Erfindung ein Kit zur Elongation, Amplifikation, Reversen Transkription, Markierung bzw. Sequenzierung von Nukleinsäuren, wobei zusätzlich Deoxynukleotide bzw. deren Derivate enthaltend sind.

Ein bevorzugter erfindungsgemäßer Reagenzienkit enthält daher zur Amplifikation von Nukleinsäuren neben den Substanzen a) oder b), welche 5'-3'-Polymeraseaktivität aufweisen, auch Deoxynukleotide oder/und Derivate davon. Gegebenenfalls können hier auch Ribonukleotide oder Derivate davon eingesetzt werden, nämlich dann, wenn eine Polymerase eingesetzt wird, welche auch Ribonukleotide als Substrat akzeptiert.

Ein weiterer bevorzugter Gegenstand der vorliegenden Erfindung ist ein Kit zur Sequenzierung von Nukleinsäuren, wobei zusätzlich zu Deoxynukleotiden bzw. deren Derivaten, Dideoxynukleotide bzw. deren Derivate zur Kettentermination enthaltend sind.

# DE 198 40 771 A 1

Desweiteren ist insbesondere Gegenstand der vorliegenden Erfindung ein Kit zur reversen Transkription von Nukleinsäuren, wobei entweder der erfundungsgemäße Komplex selbst reverse Transkriptase Aktivität besitzt, oder aber zusätzlich ein geeignetes Enzym anwesend ist, das reverse Transkriptase Aktivität besitzt, wobei Deoxynukleotide bzw. deren Derivaten im Reaktionsgemisch enthalten sind.

In einer weiteren besonders bevorzugten Ausführungsform enthält der Kit Primer oder/und Deoxynukleotide oder/und Dideoxynukleotide oder/und Ribonukleotide oder/und deren jeweilige Derivate in markierter Form. 5

Insbesondere für die Sequenzierung von Nukleinsäuren ist es nötig, eine Markierung einzufügen. Geeignete Markierungen sind weiter oben bereits in beispielhafter Form beschrieben und gelten ebenso als bevorzugte Ausführungsformen im Rahmen der erfundungsgemäßen Reagenzienkits.

Im Rahmen der vorliegenden Erfindung ist es weiterhin möglich, daß der Reagenzienkit zur Markierung von Nukleinsäuren im Rahmen einer sogenannten Nick-Translation eingesetzt wird. In diesem Fall enthält der Reagenzienkit die Bestandteile a) oder b) und markierte Nukleotide, wobei Puffersubstanzen, ATP oder andere Kofaktoren oder/und Pyrophosphat ebenfalls anwesend sein können. Primer werden allerdings im Rahmen einer Nick-Translation in der Regel nicht benötigt. 10

Gegenstand der vorliegenden Erfindung ist desweiteren die Verwendung eines thermostabilen Gleitklammerproteins in in vitro Verfahren zur Elongation, Amplifikation, Markierung bzw. Sequenzierung oder Reversen Transkription von Nukleinsäuren. 15

Bevorzugt ist der Kit, wenn als Gleitklammerprotein ein thermostabiles Homolog zu dem PCNA-Protein (SEQ ID NO: 12, 13, 14, 15), dem PCNA-Homologen aus *Archeoglobus fulgidus* oder dem  $\beta$ -clamp-Protein (SEQ ID NO: 35, 20 36) vorliegt. Insbesondere ist ein Kit bevorzugt, der zusätzlich ein Kopplungsprotein enthält.

Desweiteren ist ein Kit bevorzugt, der zusätzlich ein weiteres Protein enthaltend, wobei das weitere Protein als Klammerlader zur Verfügung steht, auch aus der Gruppe der Prokaryonten stammt und thermostabil ist. 25

Insbesondere ist ein Kit bevorzugt der einen geeigneten Puffer enthält, wie oben beschrieben. Bevorzugt ist ebenfalls, daß der erfundungsgemäße Kit zusätzlich eine Pyrophosphatase, ATP oder/und andere Kofaktoren enthält.

Die folgenden Beispiele sollen in Verbindung mit den Abbildungen die Erfindung näher erläutern: 30

## Abbildungen

Im folgenden werden häufig Sequenznamen verwendet, unter denen die Protein- oder Nukleinsäuresequenzen in der Genbank und der EMBL Datenbank stehen. 35

### Abb. 1

Proteinsequenzen, paarweise Alignments und multiple Alignments aus Archaebakterien und den entsprechenden humanen Genen des Replikationsapparates. 35

Wobei die Annotation: <sup>1</sup> bedeutet %Identität zu dem entsprechenden humanen Gen, berechnet aus dem paarweisen Alignment (siehe Anhang) mit BLASTP 2.0.4 [Feb-24-1998] und die Annotation <sup>2</sup> %Identität zu dem entsprechenden Gen von *Archaeoglobus fulgidus*, berechnet aus dem paarweisen Alignment (siehe Anhang) mit BLASTP 2.0.4 [Feb-24-1998] und die Annotation <sup>3</sup> %Identität zu dem entsprechenden humanen Gen, berechnet aus dem paarweisen Alignment mit FASTA 3.1t02 [March, 1998]<sup>4</sup> bedeutet. Die Verfahren werden näher beschrieben in: 40  
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25: 3389-3402 und W.R. Pearson & D.J. Lipman PNAS (1988) 85: 2444-2448. Die Abbildung zeigt im Falle des Gleitklammerladers I, Gleitklammerlader II, Gleitklammer, koppelnde Untereinheit und Elongationsprotein I die Sequenzen aus den Datenbanken und ebenso deren SEQ ID Nummer, wobei die Werte in den Klammern jeweils Prozent Identität je Anzahl Aminosäuren darstellen. Im Falle des Elongationsprotein II beziehen sich die Werte auf Prozent Sequenzidentität zu der *Archaeoglobus fulgidus* Sequenz. 45

### Abb. 2

Proteinsequenzen, paarweise Alignments und multiple Alignments aus Eubakterien des Replikationsapparates. Wobei die Annotation <sup>1</sup> bedeutet %Identität zu dem entsprechenden Gen aus *E. coli*, berechnet aus dem paarweisen Alignment (siehe Anhang) mit BLASTP 2.0.4 [Feb-24-1998]<sup>5</sup>. Das Verfahren wird in: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25: 3389-3402 näher beschrieben. 55

### Abb. 3

Skizze des Replikationsapparates in einer möglichen Variante, wobei die Gleitkammer über eine koppelnde Untereinheit an das Elongationsprotein bindet. 60

### Abb. 4

Die Abb. 4 zeigt Alignments zweier konservierter Regionen des Gleitklammerproteins, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: PCNA human (aus SEQ ID NO: 11), die entsprechende Sequenz aus *Archaeoglobus fulgidus* (aus SEQ ID NO: 12), aus *Methanococcus janashii* (aus SEQ ID NO: 13), aus *Pyrococcus horikoshii* (aus SEQ ID NO: 14) und aus *Methanococcus thermoautothrophicus* (aus SEQ ID NO: 15). 65

# DE 198 40 771 A 1

## Abb. 5

Die Abb. 5 zeigt ein Alignment einer konservierten Regionen der koppelnden Untereinheit, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: PfuORF2, DPD2\_HUMAN, AF1790 und MJ0702. Die SEQ ID Nummern können Abb. 1 entnommen werden.

## Abb. 6

Die Abb. 6 zeigt ein Alignment einer konservierten Regionen der koppelnden Untereinheit, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AC11\_HUMAN, AF2060, MTH0241, PHBN012 und MJ1422. Die SEQ ID Nummern können Abb. 1 entnommen werden.

## Abb. 7

Die Abb. 7 zeigt ein Alignment einer konservierten Regionen des Gleitklammerlader 2, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AC15\_HUMAN, MJ0884, AF1195, MTI10240 und MTI10240. Die SEQ ID Nummern können Abb. 1 entnommen werden.

## Abb. 8

Die Abb. 8 zeigt ein Alignment einer konservierten Regionen des Elongationsproteins 1, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: DPOD\_HUMAN, MJ0885, MTH1208, PHBT047 und DPOL\_ARCFU. Die SEQ ID Nummern können Abb. 1 entnommen werden.

## Abb. 9

Die Abb. 9 zeigt ein Alignment einer konservierten Regionen des Elongationsproteins 2, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AF1722, MJ1630, PfuORF3, MTH1536 und PHBN021. Die SEQ ID Nummern können Abb. 1 entnommen werden.

## Abb. 10

Die Abb. 10 zeigt ein Alignment einer konservierten Regionen des Elongationsproteins aus Eubakterien, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: DP3A\_ECOLI: DNA Pol III, alpha subunit, Escherichia coli, BB0579: DNA Pol III, alpha subunit, Borrelia burgdorferi, DP3A\_HELPY: DNA Pol III, alpha subunit, Helicobacter pylori AA50: Aquifex aeolicus, section 50 und DP3A\_SALTY: DNA Pol III, alpha subunit, Salmonella typhimurium).

## Abb. 11

Die Abb. 11 zeigt ein Alignment einer konservierten Regionen der Gleitklammer aus Eubakterien, sowie die daraus hergeleiteten Konsensussequenzen. Folgende Gene sind gezeigt: AAPOL3B, DP3B\_ECOLI, S.TYPHIM, DP3B\_PROMI, DP3B\_PSEPU und DP3B\_STRCO (AAPOL3B: Aquifex Aeolicus sektion 93; DP3B\_ECOLI: DNA Pol III, beta chain, Escherichia coli S.TYPHIM: DNA Pol III, beta chain, Salmonella typhimurium P3B\_PROMI: DNA Pol III, beta chain, Probeus mirabilis DP3B\_PSEPU: DNA Pol III, beta chain, Pseudomonas putida DP3B\_STRCO: DNA Pol III, beta chain, Strptomyces coelicolor).

## Abb. 12

50 Multiples Alignment der Gleitklammerproteineinsequenzen zur Generierung des Hidden Markov Models.

## Abb. 13

55 Multiples Alignment der eubakteriellen Gleitklammerproteinsequenzen zur Generierung des Hidden Markov Models (AAPOL3B: Aquifex Aeolicus sektion 93; DP3B\_ECOLI: DNA Pol III, beta chain, Escherichia coli S.TYPHIM: DNA Pol III, beta chain, Salmonella typhimurium P3B\_PROMI: DNA Pol III, beta chain, Probeus mirabilis DP3B\_PSEPU: DNA Pol III, beta chain, Pseudomonas putida DP3B\_STRCO: DNA Pol III, beta chain, Strptomyces coelicolor).

## Abb. 14

60 Multiples Alignment der Gleitklammerlader 1 Proteinsequenzen zur Generierung des Hidden Markov Models.

## Abb. 15

65 Multiples Alignment der Gleitklammerlader 2 Proteinsequenzen zur Generierung des Hidden Markov Models.

Abb. 16

Multiples Alignment der Proteinsequenzen der koppelnden Untereinheiten zur Generierung des Hidden Markov Models.

5

Abb. 17

Multiples Alignment der Sequenzen der Elongationsproteine 1 zur Generierung des Hidden Markov Models.

10

Abb. 18

Multiples Alignment der Sequenzen der Elongationsproteine 2 zur Generierung des Hidden Markov Models.

15

Abb. 19

Multiples Alignment der Sequenzen der eubakteriellen Elongationsproteine zur Generierung des Hidden Markov Models. Folgende Gene sind gezeigt:

DP3A\_ECOLI: DNA Pol III, alpha subunit, Escherichia coli, BB0579: DNA Pol III, alpha subunit, *Borrelia burgdorferi*, DP3A\_HELPY: DNA Pol III, alpha subunit, *Helicobacter pylori* AA50: *Aquifex aeolicus*, section 50 und DP3A\_SALTY: DNA Pol III, alpha subunit, *Salmonella typhimurium*).

20

### Beispiel

Die DNA wird mittels der bekannten Verfahren aus dem Organismus *Archaeoglobus fulgidus* (DSM No. 4304) gereinigt. Organismen Aufzucht geschieht hierbei durch die DSM (Deutsche Sammlung für Mikroorganismen). Um die entsprechenden Gene (Gleitklammerlader I/II, Gleitklammer, Elongationsproteine I/II, koppelnde Untereinheit) in den Expressionsvektor pTrc99 zu klonieren, werden für jedes Gen Primer entwickelt, die den vollständigen offenen Leserahmen umspannen, sowie Start- wie Stopcodon. Hierbei werden den Primern jeweils Restriktionsenden hinzugefügt, die die gerichtete Klonierung in den Expressionsvektor erleichtern. Unter Verwendung von etwa 200 ng Gesamtgenomischer DNA werden mit den entsprechenden Annealing-Temperaturen PCR Reaktionen gefahren (etwa 35 Zyklen) und die daraus resultierenden Produkte gereinigt. Nach der Reinigung werden die Produkte mit Restriktionsenzymen behandelt und über ein Agarosegel gereinigt um zur Ligation bereit zu stehen. Der Expressionsvektor wird mittels Restriktionsenzymen so linearisiert, gereinigt und verdünnt, daß er zur Ligation mit den Amplifikaten der akzessorischen Gene aus der obigen PCR bereit ist. Die Ligation wird angesetzt und nach Inkubation ein Aliquot in den *E. coli* Stamm INVAlpha<sup>+</sup> (Invitrogen) transformiert. Von jedem Gen werden 3 positive Kolonien gepickt, Plasmid DNA präpariert und die Inserts auf Vollständigkeit und Richtigkeit mittels DNA Sequenzierung überprüft. Korrekte Klone werden ausgesucht und erneut zur Vereinzelung auf Agarplatten (ampicillin) ausgebracht. Kolonien werden gepickt und Übernachtkulturen angesetzt. Ein Aliquot (500 µl) der Übernachtkultur wird in eine ein bis fünf Liter Kultur von LB (Ampicillin: 80 mg/l) gegeben. Die Kulturen wachsen bis zu einer OD<sub>660</sub> von 0.8 bei 37°C nun wird zur Induktion IPTG zugegeben (125 mg/l). Diese Kulturen wachsen nun weitere 11 Stunden. Die Kulturen werden zentrifugiert und die Pellets in einem Puffer aufgenommen (Puffer A: 50 mM Tris-HCl pH 7.9, 50 mM Dextrose, 1 mM EDTA). Nach Zentrifugation werden die Zellen erneut aufgenommen, jedoch enthält Puffer A nun zusätzlich Lysozym (4 mg/ml). Nach Inkubation (15 min.) wird ein gleiches Volumen Puffer B zugegeben (B: 10 mM Tris-HCl pH 7.9, 50 mM KCl, 1 mM EDTA, 1 mM PMSF, 0.5% Tween 20, 0.5% Nonidet P40) und die Lyse zur Inkubation bei 75°C für eine Stunde gegeben. Nach Zentrifugation wird der Überstand entnommen und die überexprimierten Proteine werden mittels (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ausgefällt. Die Pellets werden nach Zentrifugation gesammelt und die Proteine mit Puffer A resuspendiert. Die resuspendierten Proteine werden gegen Aufbewahrungspuffer (50 mM Tris-HCl pH 7.9, 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 50% Glycerol) dialysiert und anschließend bei -70°C gelagert.

25

Um die Aktivität der Proteine zu testen werden Reaktionen wie folgt zusammengesetzt: Aliquote der Proteine werden in unterschiedlichen Konfigurationen und Molaritäten vorliegen, Gleitklammerlader I/II mit Gleitklammer, koppelnder Untereinheit und Elongationsprotein I, oder Gleitklammerlader I/II mit Gleitklammer, mit und ohne koppelnder Untereinheit und Elongationsprotein II, oder Gleitklammer, und Elongationsprotein I oder II und schließlich nur Elongationsprotein I oder II; als Puffer dient der obige Aufbewahrungspuffer. DNA Polymerisationsaktivität wird mittels Inkorporation von (Methyl-<sup>3</sup>H) TTP in Trichlorsäure-unlösliches Material gemessen (Ishino, Y., Iwasaki, H., Fukui, H., Mineno, J., Kato, I., & Schinigawa, H. (1992) Biochimie 74, 131–136). Um die Prozessivität zu bestimmen werden die obigen Proteingemische in Primer-Elongationsexperimenten verwendet. Ein M13 Einzelstrangtemplate wird in 10 mM Tris-HCl (pH 9.4) eingebracht und gemeinsam mit einem Universal-primer (5'-FITC markiert) erhitzt (92°C) und abgekühlt (Raumtemperatur). Verdünnungsreihen des so generierten Template-Primer Gemisches werden in einer Reaktion bestehend aus Nukleotiden (etwa 200 µM bis 1 mM), Reaktionspuffer (Endkonzentration: 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5–5 mM MgCl<sub>2</sub>, ATP (0 mM–200 mM) und proteinstabilisierenden Agentien zusammengeführt und für 10 Minuten bei 37°C, 52°, 62°, 68°, 74°, und 78° inkubiert. Ein Aliquot wird zur Analyse auf einem Sequenzautomaten geladen (z. B. ALF, Pharmacie Biotech). Die obigen Proteingemische dienen auch dazu Fidelity und Exonukleaseaktivität zu messen, wobei die in Kohler et al. (Proc. Natl. Acad. Sci. USA 88: 7958–7962 (1991) oder Chase et al. (J. Biol. Chem., 249: 4545–4552 (1972) beschriebenen Verfahren zur Anwendung kommen. Ebenso werden die Proteingemische in der PCR eingesetzt (Methods in Molecular Biology Vol. 15 Humana Press Totowa, New Jersey, 1993, edited by Bruce A. White).

30

35

40

45

50

55

60

65

## (1) ALLGEMEINE ANGABEN:

## 5 (i) ANMELDER:

- (A) NAME: LION bioscience AG
- (B) STRASSE: Im Neuenheimer Feld 517
- (C) ORT: Heidelberg
- 10 (E) LAND: DE
- (F) POSTLEITZAHL: 69120

15 (ii) BEZEICHNUNG DER ERFINDUNG: Akzessorische Komplexe mit  
Polymeraseaktivitaet

## 20 (iii) ANZAHL DER SEQUENZEN: 38

## 25 (iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- 25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

## 30 (2) ANGABEN ZU SEQ ID NO: 1:

## 35 (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 340 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: beides

## 40 (ii) ART DES MOLEKÜLS: Protein

## 45 (vi) URSPRUNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

## 50 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

Met Glu Thr Ser Ala Leu Lys Gln Gln Glu Gln Pro Ala Ala Thr Lys  
1 5 10 15

55 Ile Arg Asn Leu Pro Trp Val Glu Lys Tyr Arg Pro Gln Thr Leu Asn  
20 25 30

60

65

## DE 198 40 771 A 1

Asp	Leu	Ile	Ser	His	Gln	Asp	Ile	Leu	Ser	Thr	Ile	Gln	Lys	Phe	Ile		
35		40		45													5
Asn	Glu	Asp	Arg	Leu	Pro	His	Leu	Leu	Leu	Tyr	Gly	Pro	Pro	Gly	Thr		
50		55		60													10
Gly	Lys	Thr	Ser	Thr	Ile	Leu	Ala	Cys	Ala	Lys	Gln	Leu	Tyr	Lys	Asp		
65		70		75		75		80									15
Lys	Glu	Phe	Gly	Ser	Met	Val	Leu	Glu	Leu	Asn	Ala	Ser	Asp	Asp	Arg		
85		90		95													20
Gly	Ile	Asp	Ile	Ile	Arg	Gly	Pro	Ile	Leu	Ser	Phe	Ala	Ser	Thr	Arg		
100		105		110													25
Thr	Ile	Phe	Lys	Lys	Gly	Phe	Lys	Leu	Val	Ile	Leu	Asp	Glu	Ala	Asp		
115		120		125													30
Ala	Met	Thr	Gln	Asp	Ala	Gln	Asn	Ala	Leu	Arg	Arg	Val	Ile	Glu	Lys		
130		135		140													35
Phe	Thr	Glu	Asn	Thr	Arg	Phe	Cys	Leu	Ile	Cys	Asn	Tyr	Leu	Ser	Lys		
145		150		155		155		160									40
Ile	Ile	Pro	Ala	Leu	Gln	Ser	Arg	Cys	Thr	Arg	Phe	Arg	Phe	Gly	Pro		
165		170		175													45
Leu	Thr	Pro	Glu	Leu	Met	Val	Pro	Arg	Leu	Glu	His	Val	Val	Glu	Glu		
180		185		190													50
Glu	Lys	Val	Asp	Ile	Ser	Glu	Asp	Gly	Met	Lys	Ala	Leu	Val	Thr	Leu		
195		200		205													55
Ser	Ser	Gly	Asp	Met	Arg	Arg	Ala	Leu	Asn	Ile	Leu	Gln	Ser	Thr	Asn		
210		215		220													60
Met	Ala	Phe	Gly	Lys	Val	Thr	Glu	Glu	Thr	Val	Tyr	Thr	Cys	Thr	Gly		
225		230		235		235		240									65
His	Pro	Leu	Lys	Ser	Asp	Ile	Ala	Asn	Ile	Leu	Asp	Trp	Met	Leu	Asn		
245		250		255													70
Gln	Asp	Phe	Thr	Thr	Ala	Tyr	Arg	Asn	Ile	Thr	Glu	Leu	Lys	Thr	Leu		
260		265		270													75
Lys	Gly	Leu	Ala	Leu	His	Asp	Ile	Leu	Thr	Glu	Ile	His	Leu	Phe	Val		
275		280		285													80

His Arg Val Asp Phe Pro Ser Ser Val Arg Ile His Leu Leu Thr Lys  
 290 295 300

5 Met Ala Asp Ile Glu Tyr Arg Leu Ser Val Gly Thr Asn Glu Lys Ile  
 305 310 315 320

10 Gln Leu Ser Ser Leu Ile Ala Ala Phe Gln Val Thr Arg Asp Leu Ile  
 325 330 335

15 Val Ala Glu Ala  
 340

## (2) ANGABEN ZU SEQ ID NO: 2:

## 20 (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 319 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

## (ii) ART DES MOLEKÜLS: Protein

30 (vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: Archaeoglobus fulgidus

## 35 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

40 Met Glu Asn Phe Glu Ile Trp Val Glu Lys Tyr Arg Pro Arg Thr Leu  
 1 5 10 15

45 Asp Glu Val Val Gly Gln Asp Glu Val Ile Gln Arg Leu Lys Gly Tyr  
 20 25 30

50 Val Glu Arg Lys Asn Ile Pro His Leu Leu Phe Ser Gly Pro Pro Gly  
 35 40 45

55 Thr Gly Lys Thr Ala Thr Ala Ile Ala Leu Ala Arg Asp Leu Phe Gly  
 50 55 60

60 Glu Asn Trp Arg Asp Asn Phe Ile Glu Met Asn Ala Ser Asp Glu Arg  
 65 70 75 80

65 Gly Ile Asp Val Val Arg His Lys Ile Lys Glu Phe Ala Arg Thr Ala  
 85 90 95

DE 198 40 771 A 1

Pro Ile Gly Gly Ala Pro Phe Lys Ile Ile Phe Leu Asp Glu Ala Asp				
100	105	110		
Ala Leu Thr Ala Asp Ala Gln Ala Ala Leu Arg Arg Thr Met Glu Met				5
115	120	125		
Tyr Ser Lys Ser Cys Arg Phe Ile Leu Ser Cys Asn Tyr Val Ser Arg				10
130	135	140		
Ile Ile Glu Pro Ile Gln Ser Arg Cys Ala Val Phe Arg Phe Lys Pro				
145	150	155	160	15
Val Pro Lys Glu Ala Met Lys Lys Arg Leu Leu Glu Ile Cys Glu Lys				
165	170	175		20
Glu Gly Val Lys Ile Thr Glu Asp Gly Leu Glu Ala Leu Ile Tyr Ile				
180	185	190		
Ser Gly Gly Asp Phe Arg Lys Ala Ile Asn Ala Leu Gln Gly Ala Ala				25
195	200	205		
Ala Ile Gly Glu Val Val Asp Ala Asp Thr Ile Tyr Gln Ile Thr Ala				
210	215	220		30
Thr Ala Arg Pro Glu Glu Met Thr Glu Leu Ile Gln Thr Ala Leu Lys				
225	230	235	240	35
Gly Asn Phe Met Glu Ala Arg Glu Leu Leu Asp Arg Leu Met Val Glu				
245	250	255		40
Tyr Gly Met Ser Gly Glu Asp Ile Val Ala Gln Leu Phe Arg Glu Ile				
260	265	270		
Ile Ser Met Pro Ile Lys Asp Ser Leu Lys Val Gln Leu Ile Asp Lys				45
275	280	285		
Leu Gly Glu Val Asp Phe Arg Leu Thr Glu Gly Ala Asn Glu Arg Ile				
290	295	300		50
Gln Leu Asp Ala Tyr Leu Ala Tyr Leu Ser Thr Leu Ala Lys Lys				
305	310	315		55

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1847 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang

65

(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Methanococcus jannaschii

10

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

Met Val Ile Ile Met Glu Lys Pro Trp Val Glu Lys Tyr Arg Pro Lys  
1 5 10 15

20

Thr Leu Asp Asp Ile Val Gly Gln Asp Glu Ile Val Lys Arg Leu Lys  
20 25 30

25

Lys Tyr Val Glu Lys Lys Ser Met Pro His Leu Leu Phe Ser Gly Pro  
35 40 4530 Pro Gly Val Gly Lys Cys Leu Thr Gly Asp Thr Lys Val Ile Val Asn  
50 55 6035 Gly Glu Ile Arg Glu Ile Gly Glu Val Ile Glu Glu Ile Ser Asn Gly  
65 70 75 80

40

Lys Phe Gly Val Thr Leu Thr Asn Asn Leu Lys Val Leu Gly Ile Asp  
85 90 95

45

Glu Asp Gly Lys Ile Arg Glu Phe Asp Val Gln Tyr Val Tyr Lys Asp  
100 105 110

50

Lys Thr Asn Thr Leu Ile Lys Ile Lys Thr Lys Met Gly Arg Glu Leu  
115 120 125

55

Lys Val Thr Thr Tyr His Pro Leu Leu Ile Asn His Lys Asn Gly Glu  
130 135 140

60

Ile Lys Trp Glu Lys Ala Glu Asn Leu Lys Val Gly Asp Lys Leu Ala  
145 150 155 160

65

Thr Pro Arg Tyr Ile Leu Phe Asn Glu Ser Asp Tyr Asn Glu Glu Leu  
165 170 175

70

Ala Glu Trp Leu Gly Tyr Phe Ile Gly Asp Gly His Ala Asp Lys Glu  
180 185 190

65

DE 198 40 771 A 1

Ser Asn Lys Ile Thr Phe Thr Asn Gly Asp Glu Lys Leu Arg Lys Arg			
195	200	205	
Phe Ala Glu Leu Thr Glu Lys Leu Phe Lys Asp Ala Lys Ile Lys Glu			5
210	215	220	
Arg Ile His Lys Asp Arg Thr Pro Asp Ile Tyr Val Asn Ser Lys Glu			10
225	230	235	240
Ala Val Glu Phe Ile Asp Lys Leu Gly Leu Arg Gly Lys Lys Ala Asp			
245	250	255	15
Lys Val Arg Ile Pro Lys Glu Ile Met Arg Ser Asp Ala Leu Arg Ala			
260	265	270	20
Phe Leu Arg Ala Tyr Phe Asp Cys Asp Gly Gly Ile Glu Lys His Ser			
275	280	285	
Ile Val Leu Ser Thr Ala Ser Lys Glu Met Ala Glu Asp Leu Val Tyr			25
290	295	300	
Ala Leu Leu Arg Phe Gly Ile Ile Ala Lys Leu Arg Glu Lys Val Asn			
305	310	315	30
Lys Asn Asn Asn Lys Val Tyr Tyr His Ile Val Ile Ser Asn Ser Ser			
325	330	335	35
Asn Leu Arg Thr Phe Leu Asp Asn Ile Gly Phe Ser Gln Glu Arg Lys			
340	345	350	
Leu Lys Lys Leu Leu Glu Ile Ile Lys Asp Glu Asn Pro Asn Leu Asp			40
355	360	365	
Val Ile Thr Ile Asp Lys Glu Lys Ile Arg Tyr Ile Arg Asp Arg Leu			45
370	375	380	
Lys Val Lys Leu Thr Arg Asp Ile Glu Lys Asp Asn Trp Ser Tyr Asn			
385	390	395	50
Lys Cys Arg Lys Ile Thr Gln Glu Leu Leu Lys Glu Ile Tyr Tyr Arg			
405	410	415	55
Leu Glu Glu Leu Lys Glu Ile Glu Lys Ala Leu Glu Glu Asn Ile Leu			
420	425	430	
Ile Asp Trp Asp Glu Val Ala Glu Arg Arg Lys Glu Ile Ala Glu Lys			60
435	440	445	

DE 198 40 771 A 1

Thr Gly Ile Arg Ser Asp Arg Arg Ile Leu Glu Tyr Ile Arg Gly Lys Arg  
450 455 460

5 Lys Pro Ser Leu Lys Asn Tyr Ile Lys Ile Ala Asn Thr Leu Gly Lys  
465 470 475 480

10 Asn Ile Glu Lys Ile Ile Asp Ala Met Arg Ile Phe Ala Lys Lys Tyr  
485 490 495

15 Ser Ser Tyr Ala Glu Ile Gly Lys Met Leu Asn Met Trp Asn Ser Ser  
500 505 510

20 Ile Lys Ile Tyr Leu Glu Ser Asn Thr Gln Glu Ile Glu Lys Leu Glu  
515 520 525

25 Glu Ile Arg Lys Thr Glu Leu Lys Leu Val Lys Glu Ile Leu Asn Asp  
530 535 540

30 Glu Lys Leu Ile Asp Ser Ile Gly Tyr Val Leu Phe Leu Ala Ser Asn  
545 550 555 560

35 Glu Ile Tyr Trp Asp Glu Ile Val Glu Ile Glu Gln Leu Asn Gly Glu  
565 570 575

40 Phe Thr Ile Tyr Asp Leu His Val Pro Arg Tyr His Asn Phe Ile Gly  
580 585 590

45 Gly Asn Leu Pro Thr Ile Leu His Asn Thr Thr Ala Ala Leu Cys Leu  
595 600 605

50 Ala Arg Asp Leu Phe Gly Glu Asn Trp Arg Asp Asn Phe Leu Glu Leu  
610 615 620

55 Asn Ala Ser Val Ser Lys Asp Thr Pro Ile Leu Val Lys Ile Asp Gly  
625 630 635 640

60 Lys Val Lys Arg Thr Thr Phe Glu Glu Leu Asp Lys Ile Tyr Phe Glu  
645 650 655

65 Thr Asn Asp Glu Asn Glu Met Tyr Lys Val Asp Asn Leu Glu Val  
660 665 670

70 Leu Thr Val Asp Glu Asn Phe Arg Val Arg Trp Arg Lys Val Ser Thr  
675 680 685

75 Ile Ile Arg His Lys Val Asp Lys Ile Leu Arg Ile Lys Phe Glu Gly  
690 695 700

## DE 198 40 771 A 1

Gly Tyr Ile Glu Leu Thr Gly Asn His Ser Ile Met Met Leu Asp Glu			
705	710	715	720
Asn Gly Leu Val Ala Lys Lys Ala Ser Asp Ile Lys Val Gly Asp Cys			
725	730	735	
Phe Leu Ser Phe Val Ala Asn Ile Glu Gly Glu Lys Asp Arg Leu Asp			
740	745	750	
Leu Lys Glu Phe Glu Pro Lys Asp Ile Thr Ser Arg Val Lys Ile Ile			
755	760	765	
Asn Asp Phe Asp Ile Asp Glu Asp Thr Ala Trp Met Leu Gly Leu Tyr			
770	775	780	
Val Ala Glu Gly Ala Val Gly Phe Lys Gly Lys Thr Ser Gly Gln Val			
785	790	795	800
Ile Tyr Thr Leu Gly Ser His Glu His Asp Leu Ile Asn Lys Leu Asn			
805	810	815	
Asp Ile Val Asp Lys Lys Gly Phe Ser Lys Tyr Glu Asn Phe Thr Gly			
820	825	830	
Ser Gly Phe Asp Arg Lys Arg Leu Ser Ala Lys Gln Ile Arg Ile Leu			
835	840	845	
Asn Thr Gln Leu Ala Arg Phe Val Glu Glu Asn Phe Tyr Asp Gly Asn			
850	855	860	
Gly Arg Arg Ala Arg Asn Lys Arg Ile Pro Asp Ile Ile Phe Glu Leu			
865	870	875	880
Lys Glu Asn Leu Arg Val Glu Phe Leu Lys Gly Leu Ala Asp Gly Asp			
885	890	895	
Ser Ser Gly Asn Trp Arg Glu Val Val Arg Ile Ser Ser Lys Ser Asp			
900	905	910	
Asn Leu Leu Ile Asp Thr Val Trp Leu Ala Arg Ile Ser Gly Ile Glu			
915	920	925	
Ser Ser Ile Phe Glu Asn Glu Ala Arg Leu Ile Trp Lys Gly Gly Met			
930	935	940	
Lys Trp Lys Lys Ser Asn Leu Leu Pro Ala Glu Pro Ile Ile Lys Met			
945	950	955	960

65

## DE 198 40 771 A 1

Ile Lys Lys Leu Glu Asn Lys Ile Asn Gly Asn Trp Arg Tyr Ile Leu  
 965 970 975

5 Arg His Gln Leu Tyr Glu Gly Lys Lys Arg Val Ser Lys Asp Lys Ile  
 980 985 990

10 Lys Gln Ile Leu Glu Met Val Asn Val Glu Lys Leu Ser Asp Lys Glu  
 995 1000 1005

15 Lys Glu Val Tyr Asp Leu Leu Lys Lys Leu Ser Lys Thr Glu Leu Tyr  
 1010 1015 1020

20 Ala Leu Val Val Lys Glu Ile Glu Ile Asp Tyr Asn Asp Phe Val  
 1025 1030 1035 1040

25 Tyr Asp Val Ser Val Pro Asn Asn Glu Met Phe Phe Ala Gly Asn Val  
 1045 1050 1055

30 Pro Ile Leu Leu His Asn Ser Asp Glu Arg Gly Ile Asp Val Ile Arg  
 1060 1065 1070

35 Thr Lys Val Lys Asp Phe Ala Arg Thr Lys Pro Ile Gly Asp Val Pro  
 1075 1080 1085

40 Phe Lys Ile Ile Phe Leu Asp Glu Ser Asp Ala Leu Thr Ala Asp Ala  
 1090 1095 1100

45 Gln Asn Ala Leu Arg Arg Thr Met Glu Lys Tyr Ser Asp Val Cys Arg  
 1105 1110 1115 1120

50 Phe Ile Leu Ser Cys Leu Thr Gly Asp Ala Lys Ile Thr Leu Pro Asp  
 1125 1130 1135

55 Glu Arg Glu Ile Lys Ile Glu Asp Phe Ile Lys Met Phe Glu Glu Arg  
 1140 1145 1150

60 Lys Leu Lys His Val Leu Asn Arg Asn Gly Glu Asp Leu Val Leu Ala  
 1155 1160 1165

65 Gly Val Lys Phe Asn Ser Lys Ile Val Asn His Lys Val Tyr Arg Leu  
 1170 1175 1180

70 Val Leu Glu Ser Gly Arg Glu Ile Glu Ala Thr Gly Asp His Lys Phe  
 1185 1190 1195 1200

75 Leu Thr Arg Asp Gly Trp Lys Glu Val Tyr Glu Leu Lys Glu Asp Asp  
 1205 1210 1215

## DE 198 40 771 A 1

Glu Val Leu Val Tyr Pro Ala Leu Glu Gly Val Gly Phe Glu Val Asp				
1220	1225	1230		
Glu Arg Arg Ile Ile Gly Leu Asn Glu Phe Tyr Glu Phe Leu Thr Asn				5
1235	1240	1245		
Tyr Glu Ile Lys Leu Gly Tyr Lys Pro Leu Gly Lys Ala Lys Ser Tyr				10
1250	1255	1260		
Lys Glu Leu Ile Thr Arg Asp Lys Glu Lys Ile Leu Ser Arg Val Leu				15
1265	1270	1275	1280	
Glu Leu Ser Asp Lys Tyr Ser Lys Ser Glu Ile Arg Arg Lys Ile Glu				20
1285	1290	1295		
Glu Glu Phe Gly Ile Lys Ile Ser Leu Thr Thr Ile Lys Asn Leu Ile				
1300	1305	1310		
Asn Gly Lys Ile Asp Gly Phe Ala Leu Lys Tyr Val Arg Lys Ile Lys				25
1315	1320	1325		
Glu Leu Gly Trp Asp Glu Ile Thr Tyr Asp Asp Glu Lys Ala Gly Ile				30
1330	1335	1340		
Phe Ala Arg Leu Leu Gly Phe Ile Ile Gly Asp Gly His Leu Ser Lys				35
1345	1350	1355	1360	
Ser Lys Glu Gly Arg Ile Leu Ile Thr Ala Thr Ile Asn Glu Leu Glu				
1365	1370	1375		
Gly Ile Lys Lys Asp Leu Glu Lys Leu Gly Ile Lys Ala Ser Asn Ile				40
1380	1385	1390		
Ile Glu Lys Asp Ile Glu His Lys Leu Asp Gly Arg Glu Ile Lys Gly				45
1395	1400	1405		
Lys Thr Ser Phe Ile Tyr Ile Asn Asn Lys Ala Phe Tyr Leu Leu				
1410	1415	1420		
Asn Phe Trp Gly Val Glu Ile Gly Asn Lys Thr Ile Asn Gly Tyr Asn				50
1425	1430	1435	1440	
Ile Pro Lys Trp Ile Lys Tyr Gly Asn Lys Phe Val Lys Arg Glu Phe				55
1445	1450	1455		
Leu Arg Gly Leu Phe Gly Ala Asp Gly Thr Lys Pro Tyr Ile Lys Lys				60
1460	1465	1470		

DE 198 40 771 A 1

Tyr Asn Ile Asn Gly Ile Lys Leu Gly Ile Arg Val Glu Asn Ile Ser  
 1475 1480 1485

<sup>5</sup> Lys Asp Lys Thr Leu Glu Phe Phe Glu Glu Val Lys Lys Met Leu Glu  
 1490 1495 1500

<sup>10</sup> Glu Phe Glu Val Glu Ser Tyr Ile Lys Val Ser Lys Ile Asp Asn Lys  
 1505 1510 1515 1520

<sup>15</sup> Asn Leu Thr Glu Leu Ile Val Lys Ala Asn Asn Lys Asn Tyr Leu Lys  
 1525 1530 1535

Tyr Leu Ser Arg Ile Ser Tyr Ala Tyr Glu Lys Asp Asn Phe Ala Arg  
 1540 1545 1550

<sup>20</sup> Leu Val Gly Glu Tyr Leu Arg Ile Lys Glu Ala Tyr Lys Asp Ile Ile  
 1555 1560 1565

<sup>25</sup> Leu Lys Glu Ile Ala Glu Asn Ala Leu Lys Glu Ala Asp Gly Glu Lys  
 1570 1575 1580

<sup>30</sup> Ser Leu Arg Glu Leu Ala Arg Lys Tyr Asn Val Pro Val Asp Phe Ile  
 1585 1590 1595 1600

<sup>35</sup> Ile Asn Gln Leu Lys Gly Lys Asp Ile Gly Leu Pro Arg Asn Phe Met  
 1605 1610 1615

Thr Phe Glu Glu Phe Leu Lys Glu Lys Val Val Asp Gly Lys Tyr Val  
 1620 1625 1630

<sup>40</sup> Ser Glu Arg Ile Ile Lys Lys Glu Cys Ile Gly Tyr Arg Asp Val Tyr  
 1635 1640 1645

<sup>45</sup> Asp Ile Thr Cys His Lys Asp Pro Ser Phe Ile Ala Asn Gly Phe Val  
 1650 1655 1660

<sup>50</sup> Ser His Asn Cys Asn Tyr Pro Ser Lys Ile Ile Pro Pro Ile Gln Ser  
 1665 1670 1675 1680

<sup>55</sup> Arg Cys Ala Val Phe Arg Phe Ser Pro Leu Lys Lys Glu Asp Ile Ala  
 1685 1690 1695

Lys Lys Leu Lys Glu Ile Ala Glu Lys Glu Gly Leu Asn Leu Thr Glu  
 1700 1705 1710

<sup>60</sup> Ser Gly Leu Glu Ala Ile Ile Tyr Val Ser Glu Gly Asp Met Arg Lys  
 1715 1720 1725

DE 198 40 771 A 1

Ala Ile Asn Val Leu Gln Thr Ala Ala Ala Leu Ser Asp Val Ile Asp					
1730	1735	1740			5
Asp Glu Ile Val Tyr Lys Val Ser Ser Arg Ala Arg Pro Glu Glu Val					
1745	1750	1755	1760		
Lys Lys Met Met Glu Leu Ala Leu Asp Gly Lys Phe Met Glu Ala Arg					10
1765	1770	1775			
Asp Leu Leu Tyr Lys Leu Met Val Glu Trp Gly Met Ser Gly Glu Asp					15
1780	1785	1790			
Ile Leu Asn Gln Met Phe Arg Glu Ile Asn Ser Leu Asp Ile Asp Glu					
1795	1800	1805			20
Arg Lys Lys Val Glu Leu Ala Asp Ala Ile Gly Glu Thr Asp Phe Arg					
1810	1815	1820			
Ile Val Glu Gly Ala Asn Glu Arg Ile Gln Leu Ser Ala Leu Leu Ala					25
1825	1830	1835	1840		
Lys Met Ala Leu Met Gly Arg					30
1845					
(2) ANGABEN ZU SEQ ID NO: 4:					
(i) SEQUENZKENNZEICHEN:					
(A) LÄNGE: 855 Aminosäuren					
(B) ART: Aminosäure					
(C) STRANGFORM: Einzelstrang					
(D) TOPOLOGIE: linear					
(ii) ART DES MOLEKÜLS: Protein					
(vi) URSPRÜNLICHE HERKUNFT:					
(A) ORGANISMUS: Pyrococcus horikoshii					
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:					
Met His Asn Met Glu Glu Val Arg Glu Val Lys Val Leu Glu Lys Pro					55
1	5	10	15		
Trp Val Glu Lys Tyr Arg Pro Gln Arg Leu Asp Glu Ile Val Gly Gln					60
20	25	30			

## DE 198 40 771 A 1

Glu His Ile Val Lys Arg Leu Lys His Tyr Val Lys Thr Gly Ser Met  
 35 40 45

5 Pro His Leu Leu Phe Ala Gly Pro Pro Gly Val Gly Lys Cys Leu Thr  
 50 55 60

10 Gly Asp Thr Lys Val Ile Ala Asn Gly Gln Leu Phe Glu Leu Arg Glu  
 65 70 75 80

15 Leu Val Glu Lys Ile Ser Gly Gly Lys Phe Gly Pro Thr Pro Val Lys  
 85 90 95

20 Gly Leu Lys Val Ile Gly Ile Asp Glu Asp Gly Lys Leu Arg Glu Phe  
 100 105 110

25 Glu Val Gln Tyr Val Tyr Lys Asp Lys Thr Glu Arg Leu Ile Arg Ile  
 115 120 125

30 Arg Thr Arg Leu Gly Arg Glu Leu Lys Val Thr Pro Tyr His Pro Leu  
 130 135 140

35 Leu Val Asn Arg Arg Asn Gly Glu Ile Lys Trp Val Lys Ala Glu Glu  
 145 150 155 160

40 Leu Lys Pro Gly Asp Lys Leu Ala Val Pro Arg Phe Leu Pro Ile Val  
 165 170 175

45 Thr Gly Glu Asp Pro Leu Ala Glu Trp Leu Gly Tyr Phe Leu Gly Gly  
 180 185 190

50 Gly Tyr Ala Asp Ser Lys Glu Asn Leu Ile Met Phe Thr Asn Glu Asp  
 195 200 205

55 Pro Leu Leu Arg Gln Arg Phe Met Glu Leu Thr Glu Lys Leu Phe Ser  
 210 215 220

60 Asp Ala Arg Ile Arg Glu Ile Thr His Glu Asn Gly Thr Ser Lys Val  
 225 230 235 240

65 Tyr Val Asn Ser Lys Lys Ala Leu Lys Leu Val Asn Ser Leu Gly Asn  
 245 250 255

70 Ala His Ile Pro Lys Glu Cys Trp Arg Gly Ile Arg Ser Phe Leu Arg  
 260 265 270

75 Ala Tyr Phe Asp Cys Asn Gly Gly Val Lys Gly Asn Ala Ile Val Leu  
 275 280 285

## DE 198 40 771 A 1

Ala Thr Ala Ser Lys Glu Met Ser Gln Glu Ile Ala Tyr Ala Leu Ala			
290	295	300	
Gly Phe Gly Ile Ile Ser Arg Ile Gln Glu Tyr Arg Val Ile Ile Ser			5
305	310	315	320
Gly Ser Asp Asn Val Lys Lys Phe Leu Asn Glu Ile Gly Phe Ile Asn			10
325	330	335	
Arg Asn Lys Leu Glu Lys Ala Leu Lys Leu Val Lys Lys Asp Asp Pro			
340	345	350	15
Gly His Asp Gly Leu Glu Ile Asn Tyr Glu Leu Ile Ser Tyr Val Lys			
355	360	365	20
Asp Arg Leu Arg Leu Ser Phe Phe Asn Asp Lys Arg Ser Trp Ser Tyr			
370	375	380	
Arg Glu Ala Lys Glu Ile Ser Trp Glu Leu Met Lys Glu Ile Tyr Tyr			25
385	390	395	400
Arg Leu Asp Glu Leu Glu Lys Leu Lys Glu Ser Leu Ser Arg Gly Ile			30
405	410	415	
Leu Ile Asp Trp Asn Glu Val Ala Lys Arg Ile Glu Glu Val Ala Glu			
420	425	430	35
Glu Thr Gly Ile Arg Ala Asp Glu Leu Leu Glu Tyr Ile Glu Gly Lys			
435	440	445	40
Arg Lys Leu Ser Phe Lys Asp Tyr Ile Lys Ile Ala Lys Val Leu Gly			
450	455	460	
Ile Asp Val Glu His Thr Ile Glu Ala Met Arg Val Phe Ala Arg Lys			45
465	470	475	480
Tyr Ser Ser Tyr Ala Glu Ile Gly Arg Arg Leu Gly Thr Trp Asn Ser			
485	490	495	50
Ser Val Lys Thr Ile Leu Glu Ser Asn Ala Val Asn Val Glu Ile Leu			
500	505	510	55
Glu Arg Ile Arg Lys Ile Glu Leu Glu Leu Ile Glu Glu Ile Leu Ser			
515	520	525	
Asp Glu Lys Leu Lys Glu Gly Ile Ala Tyr Leu Ile Phe Leu Ser Gln			60
530	535	540	

DE 198 40 771 A 1

Asn Glu Leu Tyr Trp Asp Glu Ile Thr Lys Val Glu Glu Leu Arg Gly  
 545 550 555 560

<sup>5</sup> Glu Phe Ile Ile Tyr Asp Leu His Val Pro Gly Tyr His Asn Phe Ile  
 565 570 575

<sup>10</sup> Ala Gly Asn Met Pro Thr Val Val His Asn Thr Thr Ala Ala Leu Ala  
 580 585 590

Leu Ser Arg Glu Leu Phe Gly Glu Asn Trp Arg His Asn Phe Leu Glu  
 595 600 605

Leu Asn Ala Ser Asp Glu Arg Gly Ile Asn Val Ile Arg Glu Lys Val  
 610 615 620

<sup>20</sup> Lys Glu Phe Ala Arg Thr Lys Pro Ile Gly Gly Ala Ser Phe Lys Ile  
 625 630 635 640

<sup>25</sup> Ile Phe Leu Asp Glu Ala Asp Ala Leu Thr Gln Asp Ala Gln Gln Ala  
 645 650 655

<sup>30</sup> Leu Arg Arg Thr Met Glu Met Phe Ser Ser Asn Val Arg Phe Ile Leu  
 660 665 670

<sup>35</sup> Ser Cys Asn Tyr Ser Ser Lys Ile Ile Glu Pro Ile Gln Ser Arg Cys  
 675 680 685

Ala Ile Phe Arg Phe Arg Pro Leu Arg Asp Glu Asp Ile Ala Lys Arg  
 690 695 700

<sup>40</sup> Leu Arg Tyr Ile Ala Glu Asn Glu Gly Leu Glu Leu Thr Glu Glu Gly  
 705 710 715 720

<sup>45</sup> Leu Gln Ala Ile Leu Tyr Ile Ala Glu Gly Asp Met Arg Arg Ala Ile  
 725 730 735

<sup>50</sup> Asn Ile Leu Gln Ala Ala Ala Leu Asp Lys Lys Ile Thr Asp Glu  
 740 745 750

<sup>55</sup> Asn Val Phe Met Val Ala Ser Arg Ala Arg Pro Glu Asp Ile Arg Glu  
 755 760 765

Met Met Leu Leu Ala Leu Lys Gly Asn Phe Leu Lys Ala Arg Glu Lys  
 770 775 780

<sup>60</sup> Leu Arg Glu Ile Leu Leu Lys Gln Gly Leu Ser Gly Glu Asp Val Leu  
 785 790 795 800

DE 198 40 771 A 1

Lys Pro Val Gly Ala Pro Phe Arg Ile Ile Phe Leu Asp Glu Val Asp  
100 105 110

5 Asn Met Thr Lys Asp Ala Gln His Ala Leu Arg Arg Glu Met Glu Met  
115 120 125

10 Tyr Thr Lys Thr Ser Ser Phe Ile Leu Ser Cys Asn Tyr Ser Ser Lys  
130 135 140

Ile Ile Asp Pro Ile Gln Ser Arg Cys Ala Ile Phe Arg Phe Leu Pro  
15 145 150 155 160

Leu Lys Gly His Gln Ile Ile Lys Arg Leu Glu Tyr Ile Ala Glu Lys  
165 170 175

20 Glu Asn Leu Glu Tyr Glu Ala His Ala Leu Glu Thr Ile Val Tyr Phe  
180 185 190

25 Ala Glu Gly Asp Leu Arg Lys Ala Ile Asn Leu Leu Gln Ser Ala Ala  
195 200 205

30 Ser Leu Gly Glu Lys Ile Thr Glu Ser Ser Ile Tyr Asp Val Val Ser  
210 215 220

35 Arg Ala Arg Pro Lys Asp Val Arg Lys Met Ile Lys Thr Ile Leu Asp  
225 230 235 240

Gly Lys Phe Met Glu Ala Arg Asp Met Leu Arg Glu Ile Met Val Leu  
245 250 255

40 Gln Gly Ile Ser Gly Glu Asp Met Val Thr Gln Ile Tyr Gln Glu Leu  
260 265 270

45 Ser Arg Leu Ala Met Glu Gly Glu Val Asp Gly Asp Arg Tyr Val Gly  
275 280 285

50 Leu Ile Asp Ala Ile Gly Glu Tyr Asp Phe Arg Ile Arg Glu Gly Ala  
290 295 300

55 Asn Pro Arg Ile Gln Leu Glu Ala Leu Leu Ala Arg Phe Leu Glu His  
305 310 315 320

Ala

60

65

## (2) ANGABEN ZU SEQ ID NO: 6:

## (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1148 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

5

10

## (ii).ART DES MOLEKÜLS: Protein

## (vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Homo sapiens

15

20

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

Met Asp Ile Arg Lys Phe Phe Gly Val Ile Pro Ser Gly Lys Lys Leu  
 1 5 10 15

25

Val Ser Glu Thr Val Lys Lys Asn Glu Lys Thr Lys Ser Asp Glu Glu  
 20 25 30

30

Thr Leu Lys Ala Lys Lys Gly Ile Lys Glu Ile Lys Val Asn Ser Ser  
 35 40 45

35

Arg Lys Glu Asp Asp Phe Lys Gln Lys Gln Pro Ser Lys Lys Arg  
 50 55 60

Ile Ile Tyr Asp Ser Asp Ser Glu Ser Glu Glu Thr Leu Gln Val Lys  
 65 70 75 80

40

Asn Ala Lys Lys Pro Pro Glu Lys Leu Pro Val Ser Ser Lys Pro Gly  
 85 90 95

45

Lys Ile Ser Arg Gln Asp Pro Val Thr Tyr Ile Ser Glu Thr Asp Glu  
 100 105 110

50

Glu Asp Asp Phe Met Cys Lys Ala Ala Ser Lys Ser Lys Glu Asn  
 115 120 125

55

Gly Arg Ser Thr Asn Ser His Leu Gly Thr Ser Asn Met Lys Lys Asn  
 130 135 140

Glu Glu Asn Thr Lys Thr Lys Asn Lys Pro Leu Ser Pro Ile Lys Leu  
 145 150 155 160

60

65

DE 198 40 771 A 1

Thr Pro Thr Ser Val Leu Asp Tyr Phe Gly Thr Gly Ser Val Gln Arg  
165 170 175

<sup>5</sup> Ser Asn Lys Lys Met Val Ala Ser Lys Arg Lys Glu Leu Ser Gln Asn  
180 185 190

<sup>10</sup> Thr Asp Glu Ser Gly Leu Asn Asp Glu Ala Ile Ala Lys Gln Leu Gln  
195 200 205

Leu Asp Glu Asp Ala Glu Leu Glu Arg Gln Leu His Glu Asp Glu Glu  
15 210 215 220

Phe Ala Arg Thr Leu Ala Met Leu Asp Glu Glu Pro Lys Thr Lys Lys  
225 230 235 240

<sup>20</sup> Ala Arg Lys Asp Thr Glu Ala Gly Glu Thr Phe Ser Ser Val Gln Ala  
245 250 255

<sup>25</sup> Asn Leu Ser Lys Ala Glu Lys His Lys Tyr Pro His Lys Val Lys Thr  
260 265 270

Ala Gln Val Ser Asp Glu Arg Lys Ser Tyr Ser Pro Arg Lys Gln Ser  
30 275 280 285

Lys Tyr Glu Ser Ser Lys Glu Ser Gln Gln His Ser Lys Ser Ser Ala  
35 290 295 300

Asp Lys Ile Gly Glu Val Ser Ser Pro Lys Ala Ser Ser Lys Leu Ala  
305 310 315 320

<sup>40</sup> Ile Met Lys Arg Lys Lys Glu Ser Ser Tyr Lys Glu Ile Glu Pro Val  
325 330 335

<sup>45</sup> Ala Ser Lys Arg Lys Glu Asn Ala Ile Lys Leu Lys Gly Glu Thr Lys  
340 345 350

Thr Pro Lys Lys Thr Lys Ser Ser Pro Ala Lys Lys Glu Ser Val Ser  
50 355 360 365

Pro Glu Asp Ser Glu Lys Lys Arg Thr Asn Tyr Gln Ala Tyr Arg Ser  
55 370 375 380

Tyr Leu Asn Arg Glu Gly Pro Lys Ala Leu Gly Ser Lys Glu Ile Pro  
385 390 395 400

<sup>60</sup> Lys Gly Ala Glu Asn Cys Leu Glu Gly Leu Ile Phe Val Ile Thr Gly  
405 410 415

DE 198 40 771 A 1

Val Leu Glu Ser Ile Glu Arg Asp Glu Ala Lys Ser Leu Ile Glu Arg				
420	425	430		5
Tyr Gly Gly Lys Val Thr Gly Asn Val Ser Lys Lys Thr Asn Tyr Leu				
435	440	445		
Val Met Gly Arg Asp Ser Gly Gln Ser Lys Ser Asp Lys Ala Ala Ala				10
450	455	460		
Leu Gly Thr Lys Ile Ile Asp Glu Asp Gly Leu Leu Asn Leu Ile Arg				
465	470	475	480	15
Thr Met Pro Gly Lys Lys Ser Lys Tyr Glu Ile Ala Val Glu Thr Glu				
485	490	495		20
Met Lys Lys Glu Ser Lys Leu Glu Arg Thr Pro Gln Lys Asn Val Gln				
500	505	510		
Gly Lys Arg Lys Ile Ser Pro Ser Lys Lys Glu Ser Glu Ser Lys Lys				25
515	520	525		
Ser Arg Pro Thr Ser Lys Arg Asp Ser Leu Ala Lys Thr Ile Lys Lys				
530	535	540		30
Glu Thr Asp Val Phe Trp Lys Ser Leu Asp Phe Lys Glu Gln Val Ala				
545	550	555	560	35
Glu Glu Thr Ser Gly Asp Ser Lys Ala Arg Asn Leu Ala Asp Asp Ser				
565	570	575		
Ser Glu Asn Lys Val Glu Asn Leu Leu Trp Val Asp Lys Tyr Lys Pro				40
580	585	590		
Thr Ser Leu Lys Thr Ile Ile Gly Gln Gln Gly Asp Gln Ser Cys Ala				
595	600	605		45
Asn Lys Leu Leu Arg Trp Leu Arg Asn Trp Gln Lys Ser Ser Ser Glu				
610	615	620		50
Asp Lys Lys His Ala Ala Lys Phe Gly Lys Phe Ser Gly Lys Asp Asp				
625	630	635	640	55
Gly Ser Ser Phe Lys Ala Ala Leu Leu Ser Gly Pro Pro Gly Val Gly				
645	650	655		
Lys Thr Thr Thr Ala Ser Leu Val Cys Gln Glu Leu Gly Tyr Ser Tyr				60
660	665	670		

DE 198 40 771 A 1

Val Glu Leu Asn Ala Ser Asp Thr Arg Ser Lys Ser Ser Leu Lys Ala  
675 680 685

<sup>5</sup> Ile Val Ala Glu Ser Leu Asn Asn Thr Ser Ile Lys Gly Phe Tyr Ser  
690 695 700

<sup>10</sup> Asn Gly Ala Ala Ser Ser Val Ser Thr Lys His Ala Leu Ile Met Asp  
705 710 715 720

Glu Val Asp Gly Met Ala Gly Asn Glu Asp Arg Gly Gly Ile Gln Glu  
15 725 730 735

Leu Ile Gly Leu Ile Lys His Thr Lys Ile Pro Ile Ile Cys Met Cys  
20 740 745 750

Asn Asp Arg Asn His Pro Lys Ile Arg Ser Leu Val His Tyr Cys Phe  
755 760 765

<sup>25</sup> Asp Leu Arg Phe Gln Arg Pro Arg Val Glu Gln Ile Lys Gly Ala Met  
770 775 780

Met Ser Ile Ala Phe Lys Glu Gly Leu Lys Ile Pro Pro Pro Ala Met  
30 785 790 795 800

Asn Glu Ile Ile Leu Gly Ala Asn Gln Asp Ile Arg Gln Val Leu His  
35 805 810 815

Asn Leu Ser Met Trp Cys Ala Arg Ser Lys Ala Leu Thr Tyr Asp Gln  
40 820 825 830

Ala Lys Ala Asp Ser His Arg Ala Lys Lys Asp Ile Lys Met Gly Pro  
835 840 845

<sup>45</sup> Phe Asp Val Ala Arg Lys Val Phe Ala Ala Gly Glu Glu Thr Ala His  
850 855 860

Met Ser Leu Val Asp Lys Ser Asp Leu Phe Phe His Asp Tyr Ser Ile  
50 865 870 875 880

Ala Pro Leu Phe Val Gln Glu Asn Tyr Ile His Val Lys Pro Val Ala  
55 885 890 895

Ala Gly Gly Asp Met Lys Lys His Leu Met Leu Leu Ser Arg Ala Ala  
900 905 910

<sup>60</sup> Asp Ser Ile Cys Asp Gly Asp Leu Val Asp Ser Gln Ile Arg Ser Lys  
915 920 925

DE 198 40 771 A 1

Gln Asn Trp Ser Leu Leu Pro Ala Gln Ala Ile Tyr Ala Ser Val Leu				
930	935	940		
Pro Gly Glu Leu Met Arg Gly Tyr Met Thr Gln Phe Pro Thr Phe Pro				5
945	950	955	960	
Ser Trp Leu Gly Lys His Ser Ser Thr Gly Lys His Asp Arg Ile Val				10
965	970	975		
Gln Asp Leu Ala Leu His Met Ser Leu Arg Thr Tyr Ser Ser Lys Arg				15
980	985	990		
Thr Val Asn Met Asp Tyr Leu Ser Leu Leu Arg Asp Ala Leu Val Gln				20
995	1000	1005		
Pro Leu Thr Ser Gln Gly Val Asp Gly Val Gln Asp Val Val Ala Leu				25
1010	1015	1020		
Met Asp Thr Tyr Tyr Leu Met Lys Glu Asp Phe Glu Asn Ile Met Glu				30
1025	1030	1035	1040	
Ile Ser Ser Trp Gly Gly Lys Pro Ser Pro Phe Ser Lys Leu Asp Pro				35
1045	1050	1055		
Lys Val Lys Ala Ala Phe Thr Arg Ala Tyr Asn Lys Glu Ala His Leu				40
1060	1065	1070		
Thr Pro Tyr Ser Leu Gln Ala Ile Lys Ala Ser Arg His Ser Thr Ser				45
1075	1080	1085		
Pro Ser Leu Asp Ser Glu Tyr Asn Glu Glu Leu Asn Glu Asp Asp Ser				50
1090	1095	1100		
Gln Ser Asp Glu Lys Asp Gln Asp Ala Ile Glu Thr Asp Ala Met Ile				55
1105	1110	1115	1120	
Lys Lys Lys Thr Lys Ser Ser Lys Pro Ser Lys Pro Glu Lys Asp Lys				60
1125	1130	1135		
Glu Pro Arg Lys Gly Lys Gly Lys Ser Ser Lys Lys				65
1140	1145			

(2) ANGABEN ZU SEQ ID NO: 7:

- (i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 479 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Archaeoglobus fulgidus

10

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

15

Met Leu Trp Val Glu Lys Tyr Arg Pro Lys Thr Leu Glu Glu Val Val  
1 5 10 15

20

Ala Asp Lys Ser Ile Ile Thr Arg Val Ile Lys Trp Ala Lys Ser Trp  
20 25 30

25

Lys Arg Gly Ser Lys Pro Leu Leu Ala Gly Pro Pro Gly Val Gly  
35 40 45

30

Lys Thr Ser Leu Ala Leu Ala Asn Thr Met Gly Trp Glu Ala  
50 55 60

35

Val Glu Leu Asn Ala Ser Asp Gln Arg Ser Trp Arg Val Ile Glu Arg  
65 70 75 80

40

Ile Val Gly Glu Gly Ala Phe Asn Glu Thr Ile Ser Asp Glu Gly Glu  
85 90 95

45

Phe Leu Ser Ser Arg Ile Gly Lys Leu Lys Leu Ile Ile Leu Asp Glu  
100 105 110

50

Val Asp Asn Ile His Lys Lys Glu Asp Val Gly Gly Glu Ala Ala Leu  
115 120 125

55

Ile Arg Leu Ile Lys Arg Lys Pro Ala Gln Pro Leu Ile Leu Ile Ala  
130 135 140

60

Asn Asp Pro Tyr Lys Leu Ser Pro Glu Leu Arg Asn Leu Cys Glu Met  
145 150 155 160

65

Ile Asn Phe Lys Arg Leu Thr Lys Gln Gln Val Ala Arg Val Leu Glu  
165 170 175

66

Arg Ile Ala Leu Lys Glu Gly Ile Lys Val Asp Lys Ser Val Leu Leu  
180 185 190

65

DE 198 40 771 A 1

Lys Ile Ala Glu Asn Ala Gly Gly Asp Leu Arg Ala Ala Ile Asn Asp				
195	200	205		
Phe Gln Ala Leu Ala Glu Gly Lys Glu Glu Leu Lys Pro Glu Asp Val				5
210	215	220		
Phe Leu Thr Lys Arg Thr Gln Glu Lys Asp Ile Phe Arg Val Met Gln				10
225	230	235	240	
Met Ile Phe Lys Thr Lys Asn Pro Ala Val Tyr Asn Glu Ala Met Leu				
245	250	255		15
Leu Asp Glu Ser Pro Glu Asp Val Ile His Trp Val Asp Glu Asn Leu				
260	265	270		
Pro Leu Glu Tyr Ser Gly Val Glu Leu Val Asn Ala Tyr Glu Ala Leu				20
275	280	285		
Ser Arg Ala Asp Ile Phe Leu Gly Arg Val Arg Arg Arg Gln Phe Tyr				25
290	295	300		
Arg Leu Trp Lys Tyr Ala Ser Tyr Leu Met Thr Val Gly Val Gln Gln				30
305	310	315	320	
Met Lys Glu Glu Pro Lys Lys Gly Phe Thr Arg Tyr Arg Arg Pro Ala				35
325	330	335		
Val Trp Gln Met Leu Phe Gln Leu Arg Gln Lys Arg Glu Met Thr Arg				
340	345	350		
Lys Ile Leu Glu Lys Ile Gly Lys Tyr Ser His Leu Ser Met Arg Lys				40
355	360	365		
Ala Arg Thr Glu Met Phe Pro Val Ile Lys Leu Leu Leu Lys Glu Leu				45
370	375	380		
Asp Val Asp Lys Ala Ala Thr Ile Ala Ala Phe Tyr Glu Phe Thr Lys				50
385	390	395	400	
Glu Glu Leu Glu Phe Leu Val Gly Glu Lys Gly Asp Glu Ile Trp Lys				55
405	410	415		
Tyr Val Glu Lys His Gly Met His Arg Ile Glu Asp Glu Thr Phe Leu				
420	425	430		
Glu Ser Phe Val Lys Ala Glu Lys Glu Glu Lys Glu Ser Val Glu				60
435	440	445		
				65

Glu Val Ala Glu Glu Pro Glu Glu Glu Arg Glu Glu Pro Arg Ala  
450 455 460

5 Arg Lys Lys Ala Gly Lys Asn Leu Thr Leu Asp Ser Phe Phe Ser  
465 470 475

10 (2) ANGABEN ZU SEQ ID NO: 8:

15 (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 516 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

20 (ii) ART DES MOLEKÜLS: Protein

25 (vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Methanococcus jannaschii

30 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

35 Met Leu Ser Trp Val Glu Lys Tyr Arg Pro Lys Ser Leu Lys Asp Val  
1 5 10 15

40 Ala Gly His Glu Lys Val Lys Glu Lys Leu Lys Thr Trp Ile Glu Ser  
20 25 30

45 Tyr Leu Lys Gly Glu Thr Pro Lys Pro Ile Leu Leu Val Gly Pro Pro  
35 40 45

50 Gly Cys Gly Lys Thr Thr Leu Ala Tyr Ala Leu Ala Asn Asp Tyr Gly  
55 60

55 Phe Glu Val Ile Glu Leu Asn Ala Ser Asp Lys Arg Asn Ser Ser Ala  
65 70 75 80

60 Ile Lys Lys Val Val Gly His Ala Ala Thr Ser Ser Ser Ile Phe Gly  
85 90 95

65 Lys Lys Phe Leu Ile Val Leu Asp Glu Val Asp Gly Ile Ser Gly Lys  
100 105 110

70 Glu Asp Ala Gly Gly Val Ser Glu Leu Ile Lys Val Ile Lys Lys Ala  
115 120 125

## DE 198 40 771 A 1

Lys Asn Pro Ile Ile Leu Thr Ala Asn Asp Ala Tyr Ala Pro Ser Ile			
130	135	140	
Arg Ser Leu Leu Pro Tyr Val Glu Val Ile Gln Leu Asn Pro Val His			5
145	150	155	160
Thr Asn Ser Val Tyr Lys Val Leu Lys Lys Ile Ala Glu Lys Glu Gly			10
165	170	175	
Leu Asp Val Asp Asp Lys Thr Leu Lys Met Ile Ala Gln His Ser Ala			
180	185	190	
Gly Asp Leu Arg Ser Ala Ile Asn Asp Leu Glu Ala Leu Ala Leu Ser			
195	200	205	
Gly Asp Leu Ser Tyr Glu Ala Ala Gln Lys Leu Pro Asp Arg Lys Arg			
210	215	220	
Glu Ala Asn Ile Phe Asp Ala Leu Arg Val Ile Leu Lys Thr Thr His			25
225	230	235	240
Tyr Gly Ile Ala Thr Thr Ala Leu Met Asn Val Asp Glu Thr Pro Asp			
245	250	255	
Val Val Ile Glu Trp Ile Ala Glu Asn Val Pro Lys Glu Tyr Glu Lys			
260	265	270	
Pro Glu Glu Val Ala Arg Ala Phe Glu Tyr Leu Ser Lys Ala Asp Arg			
275	280	285	
Tyr Leu Gly Arg Val Met Arg Arg Gln Asn Tyr Ser Phe Trp Lys Tyr			40
290	295	300	
Ala Thr Thr Leu Met Thr Ala Gly Val Ala Leu Ser Lys Asp Glu Lys			
305	310	315	320
Tyr Arg Lys Trp Thr Pro Tyr Ser Tyr Pro Lys Ile Phe Arg Leu Leu			
325	330	335	
Thr Lys Thr Lys Ala Glu Arg Glu Ile Leu Asn Lys Ile Leu Lys Lys			
340	345	350	
Ile Gly Glu Lys Thr His Thr Ser Ser Lys Arg Ala Arg Phe Asp Leu			
355	360	365	
Gln Met Leu Lys Leu Leu Ala Lys Glu Asn Pro Ser Val Ala Ala Asp			60
370	375	380	

DE 198 40 771 A 1

Leu Val Asp Tyr Phe Glu Ile Lys Glu Asp Glu Leu Lys Val Leu Val  
385 390 395 400

5 Gly Asp Lys Leu Ala Ser Glu Ile Leu Lys Ile Leu Lys Glu Lys Lys  
405 410 415

10 Lys Leu Glu Arg Lys Lys Lys Glu Lys Lys Leu Glu Lys Glu  
420 425 430

15 Lys Lys Lys Glu Lys Ala Lys Glu Lys Gln Ser Asn Leu Ile Ile  
435 440 445

20 Gln Pro Lys Glu Ile Lys Glu Glu Val Lys Ala Glu Val Glu Lys Lys  
450 455 460

25 Glu Glu Val Lys Glu Lys Ile Val Glu Lys Pro Lys Ala Glu Glu Val  
465 470 475 480

30 Lys Glu Lys Ser Lys Thr Glu Glu Lys Glu Thr Lys Lys Asp Lys Lys  
485 490 495

35 Lys Gly Lys Lys Lys Glu Asp Lys Gly Lys Gln Leu Thr Leu Asp  
500 505 510

Ala Phe Phe Lys  
515

(2) ANGABEN ZU SEQ ID NO: 9:

40 (i) SEQUENZKENNZEICHEN:  
(A) LÄNGE: 468 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
45 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

50 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Pyrococcus horikoshii

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

60 Met Pro Asp Val Pro Trp Ile Glu Lys Tyr Arg Pro Arg Lys Leu Ser  
1 5 10 15

DE 198 40 771 A 1

Glu	Ile	Val	Asn	Gln	Glu	Gln	Ala	Leu	Glu	Lys	Val	Arg	Ala	Trp	Ile			
20					25					30								5
Glu	Ser	Trp	Leu	His	Gly	Asn	Pro	Pro	Lys	Lys	Lys	Ala	Leu	Leu	Leu			
35					40					45								
Ala	Gly	Pro	Pro	Gly	Ser	Gly	Lys	Thr	Thr	Thr	Val	Tyr	Ala	Leu	Ala			
50					55					60							10	
His	Glu	Tyr	Asn	Phe	Glu	Val	Ile	Glu	Leu	Asn	Ala	Ser	Asp	Glu	Arg			
65					70					75			80				15	
Thr	Tyr	Asn	Lys	Ile	Ala	Arg	Tyr	Val	Gln	Ala	Ala	Tyr	Thr	Met	Asp			
85					90					95							20	
Ile	Met	Gly	Lys	Arg	Arg	Lys	Ile	Ile	Phe	Leu	Asp	Glu	Ala	Asp	Asn			
100					105					110								
Ile	Glu	Pro	Ser	Gly	Ala	Pro	Glu	Ile	Ala	Lys	Leu	Ile	Asp	Lys	Ala			
115					120					125							25	
Arg	Asn	Pro	Ile	Ile	Met	Ala	Ala	Asn	His	Tyr	Trp	Glu	Val	Pro	Lys			
130					135					140							30	
Glu	Ile	Arg	Asp	Arg	Ala	Glu	Leu	Val	Glu	Tyr	Lys	Arg	Leu	Asn	Gln			
145					150					155			160				35	
Arg	Asp	Val	Ile	Ser	Ala	Leu	Val	Arg	Ile	Leu	Lys	Arg	Glu	Gly	Ile			
165					170					175							40	
Thr	Val	Pro	Lys	Glu	Ile	Leu	Thr	Glu	Ile	Ala	Lys	Arg	Ser	Ser	Gly			
180					185					190								
Asp	Leu	Arg	Ala	Ala	Ile	Asn	Asp	Leu	Gln	Thr	Ile	Val	Ala	Gly	Gly			
195					200					205							45	
Tyr	Glu	Asp	Ala	Lys	Tyr	Val	Leu	Ala	Tyr	Arg	Asp	Val	Glu	Lys	Thr			
210					215					220							50	
Val	Phe	Gln	Ser	Leu	Gly	Met	Val	Phe	Ser	Ser	Asp	Asn	Ala	Lys	Arg			
225					230					235			240				55	
Ala	Lys	Leu	Ala	Leu	Met	Asn	Leu	Asp	Met	Ser	Pro	Asp	Glu	Phe	Leu			
245					250					255								
Leu	Trp	Val	Asp	Glu	Asn	Ile	Pro	His	Met	Tyr	Leu	Lys	Pro	Glu	Glu			
260					265					270							60	

DE 198 40 771 A 1

Met Ala Arg Ala Tyr Glu Ala Ile Ser Arg Ala Asp Ile Tyr Leu Gly  
275 280 285

5 Arg Ala Gln Arg Thr Gly Asn Tyr Ser Leu Trp Lys Tyr Ala Ile Asp  
290 295 300

10 Met Met Thr Ala Gly Val Ala Val Ala Gly Thr Lys Lys Lys Gly Phe  
305 310 315 320

15 Ala Lys Phe Tyr Pro Pro Asn Thr Leu Lys Met Leu Ala Glu Ser Lys  
325 330 335

20 Glu Glu Arg Ser Ile Arg Asp Ser Ile Ile Lys Lys Ile Met Lys Glu  
340 345 350

25 Met His Met Ser Lys Leu Glu Ala Leu Glu Thr Met Lys Ile Leu Arg  
355 360 365

30 Thr Ile Phe Glu Asn Asn Leu Asp Leu Ala Ala His Phe Thr Val Phe  
370 375 380

35 Leu Glu Leu Thr Glu Lys Glu Val Glu Phe Leu Ala Gly Lys Glu Lys  
385 390 395 400

40 Ala Gly Thr Ile Trp Gly Lys Thr Leu Ser Ile Arg Arg Arg Ile Lys  
405 410 415

45 Glu Thr Glu Lys Ile Glu Glu Lys Ala Val Glu Glu Lys Val Glu Glu  
420 425 430

50 Glu Glu Ala Glu Glu Glu Glu Glu Glu Arg Lys Glu Glu Glu Lys  
435 440 445

55 Pro Lys Ala Glu Lys Lys Gly Lys Gln Val Thr Leu Phe Asp Phe  
450 455 460

60 Ile Lys Lys Asn  
465

(2) ANGABEN ZU SEQ ID NO: 10:

65 (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 479 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

65 (ii) ART DES MOLEKÜLS: Protein

## (vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: *Methanobacterium thermoautotrophicum*

5

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

Met Ser Trp Thr Glu Lys Tyr Arg Pro Gly Ser Phe Asp Glu Val Val  
 1            5            10            15

10

Gly Asn Gln Lys Val Ile Ala Glu Ile Lys Glu Trp Ile Lys Ala Trp  
 20            25            30

15

Lys Ala Gly Lys Pro Gln Lys Pro Leu Leu Leu Val Gly Pro Pro Gly  
 35            40            45

20

Thr Gly Lys Thr Thr Leu Ala His Ile Ile Gly Lys Glu Phe Ser Asp  
 50            55            60

25

Thr Leu Glu Leu Asn Ala Ser Asp Arg Arg Ser Gln Asp Ala Leu Met  
 65            70            75            80

30

Arg Ser Ala Gly Glu Ala Ser Ala Thr Arg Ser Leu Phe Asn His Asp  
 85            90            95

Leu Lys Leu Ile Ile Leu Asp Glu Val Asp Gly Ile His Gly Asn Glu  
 100            105            110

35

Asp Arg Gly Gly Val Gln Ala Ile Asn Arg Ile Ile Lys Glu Ser Arg  
 115            120            125

40

His Pro Met Val Leu Thr Ala Asn Asp Pro Tyr Ser Lys Arg Leu Gln  
 130            135            140

45

Ser Ile Lys Pro Arg Cys Arg Val Leu Asn Leu Arg Lys Val His Thr  
 145            150            155            160

Ser Ser Ile Ala Ala Ala Leu Arg Arg Ile Cys Arg Ala Glu Gly Ile  
 165            170            175

50

Glu Cys Pro Asp Asp Val Leu Arg Glu Leu Ala Lys Arg Ser Arg Gly  
 180            185            190

55

Asp Leu Arg Ser Ala Ile Asn Asp Leu Glu Ala Met Ala Glu Gly Glu  
 195            200            205

60

Glu Arg Ile Gly Glu Glu Leu Leu Lys Met Gly Glu Lys Asp Ala Thr  
 210            215            220

65

## DE 198 40 771 A 1

Ser Asn Leu Phe Asp Ala Val Arg Ala Val Leu Lys Ser Arg Asp Val

225 230 235 240

5 Ser Lys Val Arg Glu Ala Met Arg Val Asp Asp Asp Pro Thr Leu Val

245 250 255

10 Leu Glu Phe Ile Ala Glu Asn Val Pro Arg Glu Tyr Glu Lys Pro Asn

260 265 270

15 Glu Ile Ser Arg Ala Tyr Asp Met Leu Ser Arg Ala Asp Ile Phe Phe

275 280 285

20 Gly Arg Ala Val Arg Thr Arg Asn Tyr Thr Tyr Trp Arg Tyr Ala Ser

290 295 300

25 Glu Leu Met Gly Pro Gly Val Ala Leu Ala Lys Asp Lys Thr Tyr Arg

305 310 315 320

30 Thr Arg Lys Gln Arg Ser Leu Arg Asp Ser Val Ala Ala Lys Met Ala

340 345 350

35 Gly Lys Met His Ile Ser Pro Lys Val Ala Ile Ser Met Phe Pro Tyr

355 360 365

40 Met Glu Ile Leu Phe Glu Asn Asp Glu Met Ala Tyr Asp Ile Ser Glu

370 375 380

45 Ile Lys Ala Pro Lys Arg Lys Lys Thr Pro Arg Lys Ala Glu Ile Lys

405 410 415

50 Val Gly Pro Leu Tyr Ser Gln Lys Lys Asp Lys Gly Ala Asp Lys Ser

420 425 430

55 Ile Asn Asp Lys Ala Thr Asp Lys Ser Ala Lys Thr Pro Ile Lys Ser

435 440 445

60 Ser Lys Lys Asp Asp Arg Pro Arg Asp Glu Ser Ser Ser Ser Asp

450 455 460

65 Asp Lys Lys Pro Lys Glu Lys Gln Thr Ser Leu Phe Gln Phe Ser

465 470 475

## (2) ANGABEN ZU SEQ ID NO: 11:

(i) SEQUENZKENNZEICHEN:  
 (A) LÄNGE: 261 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

5

10

(ii) ART DES MOLEKÜLS: Protein

15

(vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: Homo sapiens

20

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

Met Phe Glu Ala Arg Leu Val Gln Gly Ser Ile Leu Lys Lys Val Leu  
 1 5 10 15

25

Glu Ala Leu Lys Asp Leu Ile Asn Glu Ala Cys Trp Asp Ile Ser Ser  
 20 25 30

30

Ser Gly Val Asn Leu Gln Ser Met Asp Ser Ser His Val Ser Leu Val  
 35 40 45

35

Gln Leu Thr Leu Arg Ser Glu Gly Phe Asp Thr Tyr Arg Cys Asp Arg  
 50 55 60

Asn Leu Ala Met Gly Val Asn Leu Thr Ser Met Ser Lys Ile Leu Lys  
 65 70 75 80

40

Cys Ala Gly Asn Glu Asp Ile Ile Thr Leu Arg Ala Glu Asp Asn Ala  
 85 90 95

45

Asp Thr Leu Ala Leu Val Phe Glu Ala Pro Asn Gln Glu Lys Val Ser  
 100 105 110

50

Asp Tyr Glu Met Lys Leu Met Asp Leu Asp Val Glu Gln Leu Gly Ile  
 115 120 125

55

Pro Glu Gln Glu Tyr Ser Cys Val Val Lys Met Pro Ser Gly Glu Phe  
 130 135 140

Ala Arg Ile Cys Arg Asp Leu Ser His Ile Gly Asp Ala Val Val Ile  
 145 150 155 160

60

65

DE 198 40 771 A 1

Ser Cys Ala Lys Asp Gly Val Lys Phe Ser Ala Ser Gly Glu Leu Gly  
165 170 175

5 Asn Gly Asn Ile Lys Leu Ser Gln Thr Ser Asn Val Asp Lys Glu Glu  
180 185 190

10 Glu Ala Val Thr Ile Glu Met Asn Glu Pro Val Gln Leu Thr Phe Ala  
195 200 205

15 Leu Arg Tyr Leu Asn Phe Phe Thr Lys Ala Thr Pro Leu Ser Ser Thr  
210 215 220

20 Val Thr Leu Ser Met Ser Ala Asp Val Pro Leu Val Val Glu Tyr Lys  
225 230 235 240

25 Ile Ala Asp Met Gly His Leu Lys Tyr Tyr Leu Ala Pro Lys Ile Glu  
245 250 255

30 Asp Glu Glu Gly Ser  
260

35 (2) ANGABEN ZU SEQ ID NO: 12:

(i) SEQUENZKENNZEICHEN:  
(A) LÄNGE: 245 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: Protein

45 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Archaeoglobus fulgidus

50 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

55 Met Ile Asp Val Ile Met Thr Gly Glu Leu Leu Lys Thr Val Thr Arg  
1 5 10 15

60 Ala Ile Val Ala Leu Val Ser Glu Ala Arg Ile His Phe Leu Glu Lys  
20 25 30

65 Gly Leu His Ser Arg Ala Val Asp Pro Ala Asn Val Ala Met Val Ile  
35 40 45

DE 198 40 771 A 1

Val	Asp	Ile	Pro	Lys	Asp	Ser	Phe	Glu	Val	Tyr	Asn	Ile	Asp	Glu	Glu	
50																5
Lys	Thr	Ile	Gly	Val	Asp	Met	Asp	Arg	Ile	Phe	Asp	Ile	Ser	Lys	Ser	
65																10
Ile	Ser	Thr	Lys	Asp	Leu	Val	Glu	Leu	Ile	Val	Glu	Asp	Glu	Ser	Thr	
85																15
Leu	Lys	Val	Lys	Phe	Gly	Ser	Val	Glu	Tyr	Lys	Val	Ala	Leu	Ile	Asp	
100																20
Pro	Ser	Ala	Ile	Arg	Lys	Glu	Pro	Arg	Ile	Pro	Glu	Leu	Glu	Leu	Pro	
115																25
Ala	Lys	Ile	Val	Met	Asp	Ala	Gly	Glu	Phe	Lys	Ala	Ile	Ala	Ala		
130																30
Ala	Asp	Lys	Ile	Ser	Asp	Gln	Val	Ile	Phe	Arg	Ser	Asp	Lys	Glu	Gly	
145																35
Phe	Arg	Ile	Glu	Ala	Lys	Gly	Asp	Val	Asp	Ser	Ile	Val	Phe	His	Met	
165																40
Thr	Glu	Thr	Glu	Leu	Ile	Glu	Phe	Asn	Gly	Gly	Glu	Ala	Arg	Ser	Met	
180																45
Phe	Ser	Val	Asp	Tyr	Leu	Lys	Glu	Phe	Cys	Lys	Val	Ala	Gly	Ser	Gly	
195																50
Asp	Leu	Leu	Thr	Ile	His	Leu	Gly	Thr	Asn	Tyr	Pro	Val	Arg	Leu	Val	
210																55
Phe	Glu	Leu	Val	Gly	Gly	Arg	Ala	Lys	Val	Glu	Tyr	Ile	Leu	Ala	Pro	
225																60
Arg	Ile	Glu	Ser	Glu												
245																

(2) ANGABEN ZU SEQ ID NO: 13:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 247 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: *Methanococcus jannaschii*

5

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

10 Met Phe Arg Gly Val Met Glu Ser Ala Lys Glu Phe Lys Lys Val Val  
 1 5 10 15

15 Asp Thr Ile Ser Thr Leu Leu Asp Glu Ile Cys Phe Glu Val Asp Glu  
 20 25 30

20 Glu Gly Ile Lys Ala Ser Ala Met Asp Pro Ser His Val Ala Leu Val  
 35 40 45

25 Ser Leu Glu Ile Pro Arg Leu Ala Phe Glu Glu Tyr Glu Ala Asp Ser  
 50 55 60

30 His Asp Ile Gly Ile Asp Leu Glu Ala Phe Lys Lys Val Met Asn Arg  
 65 70 75 80

35 Ala Lys Ala Lys Asp Arg Leu Ile Leu Glu Leu Asp Glu Glu Lys Asn  
 85 90 95

40 Lys Leu Asn Val Ile Phe Glu Asn Thr Gly Lys Arg Lys Phe Ser Leu  
 100 105 110

45 Ala Leu Leu Asp Ile Ser Ala Ser Ser Val Lys Val Pro Glu Ile Glu  
 115 120 125

50 Tyr Pro Asn Val Ile Met Ile Lys Gly Asp Ala Phe Lys Glu Ala Leu  
 130 135 140

55 Lys Asp Ala Asp Leu Phe Ser Asp Tyr Val Ile Leu Lys Ala Asp Glu  
 145 150 155 160

60 Asp Lys Phe Val Ile His Ala Lys Gly Asp Leu Asn Glu Asn Glu Ala  
 165 170 175

65 Ile Phe Glu Lys Asp Ser Ser Ala Ile Ile Ser Leu Glu Val Lys Glu  
 180 185 190

70 Glu Ala Lys Ser Ala Phe Asn Leu Asp Tyr Leu Met Asp Met Val Lys  
 195 200 205

75 Gly Val Ser Ser Gly Asp Ile Ile Lys Ile Tyr Leu Gly Asn Asp Met  
 210 215 220

65

Pro Leu Lys Leu Glu Tyr Ser Ile Ala Gly Val Asn Leu Thr Phe Leu  
 225                    230                    235                    240  
 Leu Ala Pro Arg Ile Glu Gly  
 245

5

## (2) ANGABEN ZU SEQ ID NO: 14:

10

## (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 249 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

15

## (ii) ART DES MOLEKÜLS: Protein

20

## (vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Pyrococcus horikoshii

25

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

30

Met Pro Phe Glu Ile Val Phe Glu Gly Ala Lys Glu Phe Ala Gln Leu  
 1                5                10                15

35

Ile Glu Thr Ala Ser Arg Leu Ile Asp Glu Ala Ala Phe Lys Val Thr  
 20                25                30

Glu Glu Gly Ile Ser Met Arg Ala Met Asp Pro Ser Arg Val Val Leu  
 35                40                45

40

Ile Asp Leu Asn Leu Pro Ser Ser Ile Phe Ser Lys Tyr Glu Val Asp  
 50                55                60

45

Gly Glu Glu Thr Ile Gly Val Asn Met Asp His Leu Lys Lys Val Leu  
 65                70                75                80

50

Lys Arg Gly Lys Ala Lys Asp Thr Leu Ile Leu Arg Lys Gly Glu Glu  
 85                90                95

55

Asn Phe Leu Glu Ile Ser Leu Gln Gly Thr Ala Thr Arg Thr Phe Arg  
 100                105                110

Leu Pro Leu Ile Asp Val Glu Glu Ile Glu Val Glu Leu Pro Asp Leu  
 115                120                125

60

65

DE 198 40 771 A 1

Pro Tyr Thr Ala Lys Val Val Val Leu Gly Val Leu Lys Glu Ala  
130 135 140

5 Val Lys Asp Ala Ser Leu Val Ser Asp Ser Ile Lys Phe Met Ala Lys  
145 150 155 160

10 Glu Asn Glu Phe Ile Met Arg Ala Glu Gly Glu Thr Gln Glu Val Glu  
165 170 175

15 Val Lys Leu Thr Leu Glu Asp Glu Gly Leu Leu Asp Ile Glu Val Gln  
180 185 190

20 Glu Glu Thr Lys Ser Ala Tyr Gly Val Ser Tyr Leu Ala Asp Met Val  
195 200 205

25 Lys Gly Ile Gly Lys Ala Asp Glu Val Thr Met Arg Phe Gly Asn Glu  
210 215 220

30 Met Pro Met Gln Met Glu Tyr Tyr Ile Arg Asp Glu Gly Arg Leu Thr  
225 230 235 240

35 Phe Leu Leu Ala Pro Arg Val Glu Glu  
245

(2) ANGABEN ZU SEQ ID NO: 15:

35 (i) SEQUENZKENNZEICHEN:  
(A) LÄNGE: 244 Aminosäuren  
(B) ART: Aminosäure  
40 (C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: Protein

45 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Methanobacterium thermoautotrophicum

50

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

55 Met Phe Lys Ala Glu Leu Asn Asp Pro Asn Ile Leu Arg Thr Ser Phe  
1 5 10 15

60 Asp Ala Ile Ser Ser Ile Val Asp Glu Val Gln Ile Gln Leu Ser Ala  
20 25 30

65

DE 198 40 771 A 1

Glu	Gly	Leu	Arg	Leu	Asp	Ala	Leu	Asp	Arg	Ser	His	Ile	Thr	Tyr	Val		5
35				40					45								
His	Leu	Glu	Leu	Lys	Ala	Glu	Leu	Phe	Asp	Glu	Tyr	Val	Cys	Asp	Glu		10
50		55			60												
Pro	Glu	Arg	Ile	Asn	Val	Asp	Thr	Glu	Glu	Leu	Met	Lys	Val	Leu	Lys		15
65				70			75			80							
Arg	Ala	Lys	Ala	Asn	Asp	Arg	Val	Ile	Leu	Ser	Thr	Asp	Glu	Gly	Asn		20
85					90				95								
Leu	Ile	Ile	Gln	Phe	Glu	Gly	Glu	Ala	Val	Arg	Thr	Phe	Lys	Ile	Arg		25
100				105					110								
Leu	Ile	Asp	Ile	Glu	Tyr	Glu	Thr	Pro	Ser	Pro	Pro	Glu	Ile	Glu	Tyr		30
115				120			125										
Glu	Asn	Glu	Phe	Glu	Val	Pro	Phe	Gln	Leu	Leu	Lys	Asp	Ser	Ile	Ala		35
130				135			140										
Asp	Ile	Asp	Ile	Phe	Ser	Asp	Lys	Ile	Thr	Phe	Arg	Val	Asp	Glu	Asp		40
145				150			155			160							
Arg	Phe	Ile	Ala	Ser	Ala	Glu	Gly	Glu	Phe	Gly	Asp	Ala	Gln	Ile	Glu		45
165				170			175										
Tyr	Leu	His	Gly	Glu	Arg	Ile	Asp	Lys	Pro	Ala	Arg	Ser	Ile	Tyr	Ser		50
180			185			190											
Leu	Asp	Lys	Ile	Lys	Glu	Met	Leu	Lys	Ala	Asp	Lys	Phe	Ser	Glu	Thr		55
195				200			205										
Ala	Ile	Ile	Asn	Leu	Gly	Asp	Asp	Met	Pro	Leu	Lys	Leu	Thr	Leu	Lys		60
210				215			220										
Met	Ala	Ser	Lys	Glu	Gly	Glu	Leu	Ser	Phe	Leu	Leu	Ala	Pro	Arg	Ile		65
225				230			235			240							
Glu	Ala	Glu	Glu														

(2) ANGABEN ZU SEQ ID NO: 16:

- (i) SEQUENZKENNZEICHEN:
  - (A) LÄNGE: 469 Aminosäuren
  - (B) ART: Aminosäure
  - (C) STRANGFORM: Einzelstrang

## (D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Homo sapiens

10

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

15

Met Phe Ser Glu Gln Ala Ala Gln Arg Ala His Thr Leu Leu Ser Pro  
1 5 10 15

20

Pro Ser Ala Asn Asn Ala Thr Phe Ala Arg Val Pro Val Ala Thr Tyr  
20 25 30

25

Thr Asn Ser Ser Gln Pro Phe Arg Leu Gly Glu Arg Ser Phe Ser Arg  
35 40 45

30

Gln Tyr Ala His Ile Tyr Ala Thr Arg Leu Ile Gln Met Arg Pro Phe  
50 55 60

35

Leu Glu Asn Arg Ala Gln Gln His Trp Gly Ser Gly Val Gly Val Lys  
65 70 75 80

40

Lys Leu Cys Glu Leu Gln Pro Glu Glu Lys Cys Cys Val Val Gly Thr  
85 90 95

45

Leu Phe Lys Ala Met Pro Leu Gln Pro Ser Ile Leu Arg Glu Val Ser  
100 105 110Glu Glu His Asn Leu Leu Pro Gln Pro Pro Arg Ser Lys Tyr Ile His  
115 120 125

50

Pro Asp Asp Glu Leu Val Leu Glu Asp Glu Leu Gln Arg Ile Lys Leu  
130 135 140Lys Gly Thr Ile Asp Val Ser Lys Leu Val Thr Gly Thr Val Leu Ala  
145 150 155 160

55

Val Phe Gly Ser Val Arg Asp Asp Gly Lys Phe Leu Val Glu Asp Tyr  
165 170 175

60

Cys Phe Ala Asp Leu Ala Pro Gln Lys Pro Ala Pro Pro Leu Asp Thr  
180 185 190

65

## DE 198 40 771 A 1

Asp Arg Phe Val Leu Leu Val Ser Gly Leu Gly Leu Gly Gly Gly				
195	200	205		
Gly Glu Ser Leu Leu Gly Thr Gln Leu Leu Val Asp Val Val Thr Gly				5
210	215	220		
Gln Leu Gly Asp Glu Gly Glu Gln Cys Ser Ala Ala His Val Ser Arg				10
225	230	235	240	
Val Ile Leu Ala Gly Asn Leu Leu Ser His Ser Thr Gln Ser Arg Asp				
245	250	255		15
Ser Ile Asn Lys Ala Lys Tyr Leu Thr Lys Lys Thr Gln Ala Ala Ser				
260	265	270		20
Val Glu Ala Val Lys Met Leu Asp Glu Ile Leu Leu Gln Leu Ser Ala				
275	280	285		
Ser Val Pro Val Asp Val Met Pro Gly Glu Phe Asp Pro Thr Asn Tyr				25
290	295	300		
Thr Leu Pro Gln Gln Pro Leu His Pro Cys Met Phe Pro Leu Ala Thr				
305	310	315	320	30
Ala Tyr Ser Thr Leu Gln Leu Val Thr Asn Pro Tyr Gln Ala Thr Ile				
325	330	335		35
Asp Gly Val Arg Phe Leu Gly Thr Ser Gly Gln Asn Val Ser Asp Ile				
340	345	350		
Phe Arg Tyr Ser Ser Met Glu Asp His Leu Glu Ile Leu Glu Trp Thr				40
355	360	365		
Leu Arg Val Arg His Ile Ser Pro Thr Ala Pro Asp Thr Leu Gly Cys				
370	375	380		45
Tyr Pro Phe Tyr Lys Thr Asp Pro Phe Ile Phe Pro Glu Cys Pro His				
385	390	395	400	50
Val Tyr Phe Cys Gly Asn Thr Pro Ser Phe Gly Ser Lys Ile Ile Arg				
405	410	415		55
Gly Pro Glu Asp Gln Thr Val Leu Leu Val Thr Val Pro Asp Phe Ser				
420	425	430		
Ala Thr Gln Thr Ala Cys Leu Val Asn Leu Arg Ser Leu Ala Cys Gln				60
435	440	445		

Pro Ile Ser Phe Ser Gly Phe Gly Ala Glu Asp Asp Asp Leu Gly Gly  
 450 455 460

5 Leu Gly Leu Gly Pro  
 465

10 (2) ANGABEN ZU SEQ ID NO: 17:

(i) SEQUENZKENNZEICHEN:  
 15 (A) LÄNGE: 488 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

20 (ii) ART DES MOLEKÜLS: Protein

25 (vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: *Archaeoglobus fulgidus*

30 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

35 Met Val Ile Lys Asn Ile Asp Ala Ala Thr Val Ala Lys Lys Phe Leu  
 1 5 10 15

40 Val Arg Gly Tyr Asn Ile Asp Pro Lys Ala Ala Glu Leu Ile Cys Lys  
 20 25 30

45 Ser Gly Leu Phe Ser Asp Glu Leu Val Asp Lys Ile Cys Arg Ile Ala  
 35 40 45

50 Asn Gly Gly Phe Ile Ile Glu Lys Ser Val Val Glu Glu Phe Leu Arg  
 50 55 60

55 Asn Leu Ser Asn Leu Lys Pro Ala Thr Leu Thr Pro Arg Pro Glu Glu  
 65 70 75 80

60 Arg Lys Val Glu Glu Val Lys Ala Ser Cys Ile Ala Leu Lys Val Ile  
 85 90 95

65 Lys Asp Ile Thr Gly Lys Ser Ser Cys Gln Gly Asn Val Glu Asp Phe  
 100 105 110

70 Leu Met Tyr Phe Asn Ser Arg Leu Glu Lys Leu Ser Arg Ile Ile Arg  
 115 120 125

DE 198 40 771 A 1

Ser Arg Val Asn Thr Thr Pro Ile Ala His Ala Gly Lys Val Arg Gly			
130	135	140	
Asn Val Ser Val Val Gly Met Val Asn Glu Val Tyr Glu Arg Gly Asp			5
145	150	155	160
Lys Cys Tyr Ile Arg Leu Glu Asp Thr Thr Gly Thr Ile Thr Cys Val			10
165	170	175	
Ala Thr Gly Lys Asn Ala Glu Val Ala Arg Glu Leu Leu Gly Asp Glu			
180	185	190	15
Val Ile Gly Val Thr Gly Leu Leu Lys Gly Ser Ser Leu Tyr Ala Asn			
195	200	205	20
Arg Ile Val Phe Pro Asp Val Pro Ile Asn Gly Asn Gly Glu Lys Lys			
210	215	220	
Arg Asp Phe Tyr Ile Val Phe Leu Ser Asp Thr His Phe Gly Ser Lys			25
225	230	235	240
Glu Phe Leu Glu Lys Glu Trp Glu Met Phe Val Arg Trp Leu Lys Gly			
245	250	255	30
Glu Val Gly Gly Lys Lys Ser Gln Asn Leu Ala Glu Lys Val Lys Tyr			
260	265	270	35
Ile Val Ile Ala Gly Asp Ile Val Asp Gly Ile Gly Val Tyr Pro Gly			
275	280	285	40
Gln Glu Asp Asp Leu Ala Ile Ser Asp Ile Tyr Gly Gln Tyr Glu Phe			
290	295	300	
Ala Ala Ser His Leu Asp Glu Ile Pro Lys Glu Ile Lys Ile Val			45
305	310	315	320
Ser Pro Gly Asn His Asp Ala Val Arg Gln Ala Glu Pro Gln Pro Ala			
325	330	335	50
Phe Glu Gly Glu Ile Arg Ser Leu Phe Pro Lys Asn Val Glu His Val			
340	345	350	55
Gly Asn Pro Ala Tyr Val Asp Ile Glu Gly Val Lys Val Leu Ile Tyr			
355	360	365	
His Gly Arg Ser Ile Asp Asp Ile Ile Ser Lys Ile Pro Arg Leu Ser			60
370	375	380	

DE 198 40 771 A 1

Tyr Asp Glu Pro Gln Lys Val Met Glu Glu Leu Leu Lys Arg Arg His  
385 390 395 400

5 Leu Ser Pro Ile Tyr Gly Gly Arg Thr Pro Leu Ala Pro Glu Arg Glu  
405 410 415

10 Asp Tyr Leu Val Ile Glu Asp Val Pro Asp Ile Leu His Cys Gly His  
420 425 430

15 Ile His Thr Tyr Gly Thr Gly Phe Tyr Arg Gly Val Phe Met Val Asn  
435 440 445

20 Ser Ser Thr Trp Gln Ala Gln Thr Glu Phe Gln Lys Lys Val Asn Leu  
450 455 460

25 Asn Pro Met Pro Gly Asn Val Ala Val Tyr Arg Pro Gly Gly Glu Val  
465 470 475 480

30 Ile Arg Leu Arg Phe Tyr Gly Glu  
485

(2) ANGABEN ZU SEQ ID NO: 18:

35 (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 594 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: Protein

45 (vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Methanococcus jannaschii

50 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

55 Met Glu Ile Ile Asn Lys Phe Leu Asp Leu Glu Ala Leu Leu Ser Pro  
1 5 10 15

60 Thr Val Tyr Glu Lys Leu Lys Asn Phe Asp Glu Glu Lys Leu Lys Arg  
20 25 30

65 Leu Ile Gln Lys Ile Arg Glu Phe Lys Lys Tyr Asn Asn Ala Phe Ile  
35 40 45

## DE 198 40 771 A 1

Leu	Leu	Asp	Glu	Lys	Phe	Leu	Asp	Ile	Phe	Leu	Gln	Lys	Asp	Leu	Asp		
50																	
Glu	Ile	Ile	Asn	Glu	Tyr	Lys	Asp	Phe	Asp	Phe	Ile	Phe	Tyr	Tyr	Thr		5
65																	
Gly	Glu	Glu	Glu	Lys	Glu	Lys	Pro	Lys	Glu	Val	Lys	Lys	Glu	Ile	Lys		10
Lys	Glu	Thr	Glu	Glu	Lys	Ile	Glu	Lys	Glu	Lys	Ile	Glu	Phe	Val	Lys		15
Lys	Glu	Glu	Glu	Gln	Phe	Ile	Lys	Lys	Ser	Asp	Glu	Asp	Val	Glu			20
Glu	Lys	Leu	Lys	Gln	Leu	Ile	Ser	Lys	Glu	Glu	Lys	Glu	Asp	Phe			
Asp	Ala	Glu	Arg	Ala	Lys	Arg	Tyr	Glu	His	Ile	Thr	Lys	Ile	Lys	Glu		25
Ser	Val	Asn	Ser	Arg	Ile	Lys	Trp	Ile	Ala	Lys	Asp	Ile	Asp	Ala	Val		30
Ile	Glu	Ile	Tyr	Glu	Asp	Ser	Asp	Val	Ser	Gly	Lys	Ser	Thr	Cys	Thr		35
Gly	Thr	Ile	Glu	Asp	Phe	Val	Lys	Tyr	Phe	Arg	Asp	Arg	Phe	Glu	Arg		
Leu	Lys	Val	Phe	Ile	Glu	Arg	Lys	Ala	Gln	Arg	Lys	Gly	Tyr	Pro	Leu		40
Lys	Asp	Ile	Lys	Lys	Met	Lys	Gly	Gln	Lys	Asp	Ile	Phe	Val	Val	Gly		45
Ile	Val	Ser	Asp	Val	Asp	Ser	Thr	Arg	Asn	Gly	Asn	Leu	Ile	Val	Arg		50
Ile	Glu	Asp	Thr	Glu	Asp	Glu	Ala	Thr	Leu	Ile	Leu	Pro	Lys	Glu	Lys		55
Ile	Glu	Ala	Gly	Lys	Ile	Pro	Asp	Asp	Ile	Leu	Leu	Asp	Glu	Val	Ile		60
Gly	Ala	Ile	Gly	Thr	Val	Ser	Lys	Ser	Gly	Ser	Ser	Ile	Tyr	Val	Asp		65

DE 198 40 771 A 1

Glu Ile Ile Arg Pro Ala Leu Pro Pro Lys Glu Pro Lys Arg Ile Asp  
 305 310 315 320

<sup>5</sup> Glu Glu Ile Tyr Met Ala Phe Leu Ser Asp Ile His Val Gly Ser Lys  
 325 330 335

<sup>10</sup> Glu Phe Leu His Lys Glu Phe Glu Lys Phe Ile Arg Phe Leu Asn Gly  
 340 345 350

<sup>15</sup> Asp Val Asp Asn Glu Leu Glu Glu Lys Val Val Ser Arg Leu Lys Tyr  
 355 360 365

<sup>20</sup> Ile Cys Ile Ala Gly Asp Leu Val Asp Gly Val Gly Val Tyr Pro Gly  
 370 375 380

<sup>25</sup> Gln Glu Glu Asp Leu Tyr Glu Val Asp Ile Ile Glu Gln Tyr Arg Glu  
 385 390 395 400

<sup>30</sup> Ile Ala Met Tyr Leu Asp Gln Ile Pro Glu His Ile Ser Ile Ile  
 405 410 415

<sup>35</sup> Ser Pro Gly Asn His Asp Ala Val Arg Pro Ala Glu Pro Gln Pro Lys  
 420 425 430

Leu Pro Glu Lys Ile Thr Lys Leu Phe Asn Arg Asp Asn Ile Tyr Phe  
 435 440 445

<sup>40</sup> Val Gly Asn Pro Cys Thr Leu Asn Ile His Gly Phe Asp Thr Leu Leu  
 450 455 460

<sup>45</sup> Tyr His Gly Arg Ser Phe Asp Asp Leu Val Gly Gln Ile Arg Ala Ala  
 465 470 475 480

<sup>50</sup> Ser Tyr Glu Asn Pro Val Thr Ile Met Lys Glu Leu Ile Lys Arg Arg  
 485 490 495

Leu Leu Cys Pro Thr Tyr Gly Gly Arg Cys Pro Ile Ala Pro Glu His  
 500 505 510

<sup>55</sup> Lys Asp Tyr Leu Val Ile Asp Arg Asp Ile Asp Ile Leu His Thr Gly  
 515 520 525

His Ile His Ile Asn Gly Tyr Gly Ile Tyr Arg Gly Val Val Met Val  
 530 535 540

<sup>60</sup> Asn Ser Gly Thr Phe Gln Glu Gln Thr Asp Phe Gln Lys Arg Met Gly  
 545 550 555 560

DE 198 40 771 A 1

Ile Ser Pro Thr Pro Ala Ile Val Pro Ile Ile Asn Met Ala Lys Val			
565	570	575	
Gly Glu Lys Gly His Tyr Leu Glu Trp Asp Arg Gly Val Leu Glu Val			5
580	585	590	
Arg Tyr			10

(2) ANGABEN ZU SEQ ID NO: 19:

(i) SEQUENZKENNZEICHEN:  
 (A) LÄNGE: 622 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: Pyrococcus horikoshii

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

Met Asp Glu Phe Val Lys Gly Leu Met Lys Asn Gly Tyr Leu Ile Thr			35
1	5	10	15

Pro Ser Ala Tyr Tyr Leu Leu Val Gly His Phe Asn Glu Gly Lys Phe			40
20	25	30	

Ser Leu Ile Glu Leu Ile Lys Phe Ala Lys Ser Arg Glu Thr Phe Ile			45
35	40	45	

Ile Asp Asp Glu Ile Ala Asn Glu Phe Leu Lys Ser Ile Gly Ala Glu			
50	55	60	

Val Glu Leu Pro Gln Glu Ile Lys Glu Gly Tyr Ile Ser Thr Gly Glu			
65	70	75	80

Gly Ser Gln Lys Val Pro Asp His Glu Glu Leu Glu Lys Ile Thr Asn			
85	90	95	

Glu Ser Ser Val Glu Ser Ser Ile Ser Thr Gly Glu Thr Pro Lys Thr			
100	105	110	

## DE 198 40 771 A 1

Glu Glu Leu Gln Pro Thr Leu Asp Ile Leu Glu Glu Glu Ile Gly Asp  
 115 120 125

5 Ile Glu Gly Gly Glu Ser Ser Ile Ser Thr Gly Asp Glu Val Pro Glu  
 130 135 140

10 Val Glu Asn Asn Asn Gly Gly Thr Val Val Val Phe Asp Lys Tyr Gly  
 145 150 155 160

15 Tyr Pro Phe Thr Tyr Val Pro Glu Glu Ile Glu Glu Leu Glu Glu  
 165 170 175

20 Tyr Pro Lys Tyr Glu Asp Val Thr Ile Glu Ile Asn Pro Asn Leu Glu  
 180 185 190

25 Val Val Pro Ile Glu Lys Asp Tyr Glu Ile Lys Phe Asp Val Arg Arg  
 195 200 205

30 Val Lys Leu Lys Pro Pro Lys Val Lys Ser Gly Ser Gly Lys Glu Gly  
 210 215 220

35 Glu Ile Ile Val Glu Ala Tyr Ala Ser Leu Phe Arg Ser Arg Leu Arg  
 225 230 235 240

Lys Leu Arg Arg Ile Leu Arg Glu Asn Pro Glu Val Ser Asn Val Ile  
 245 250 255

40 Asp Ile Lys Lys Leu Lys Tyr Val Lys Gly Asp Glu Glu Val Thr Ile  
 260 265 270

Ile Gly Leu Val Asn Ser Lys Lys Glu Thr Ser Lys Gly Leu Ile Phe  
 275 280 285

45 Glu Val Glu Asp Gln Thr Asp Arg Val Lys Val Phe Leu Pro Lys Asp  
 290 295 300

50 Ser Glu Asp Tyr Arg Glu Ala Leu Lys Val Leu Pro Asp Ala Val Val  
 305 310 315 320

Ala Phe Lys Gly Val Tyr Ser Lys Arg Gly Ile Phe Phe Ala Asn Arg  
 325 330 335

55 Phe Tyr Leu Pro Asp Val Pro Leu Tyr Arg Lys Gln Lys Pro Pro Leu  
 340 345 350

60 Glu Glu Lys Val Tyr Ala Val Leu Thr Ser Asp Ile His Val Gly Ser  
 355 360 365

DE 198 40 771 A 1

Lys Glu Phe Cys Glu Lys Ala Phe Ile Lys Phe Leu Glu Trp Leu Asn			
370	375	380	
Gly Tyr Val Glu Ser Lys Glu Glu Glu Ile Val Ser Arg Ile Arg			5
385	390	395	400
Tyr Leu Ile Ile Ala Gly Asp Val Val Asp Gly Ile Gly Ile Tyr Pro			10
405	410	415	
Gly Gln Tyr Ser Asp Leu Ile Ile Pro Asp Ile Phe Asp Gln Tyr Glu			15
420	425	430	
Ala Leu Ala Asn Leu Leu Ser Asn Val Pro Lys His Ile Thr Ile Phe			20
435	440	445	
Ile Gly Pro Gly Asn His Asp Ala Ala Arg Pro Ala Ile Pro Gln Pro			25
450	455	460	
Glu Phe Tyr Glu Glu Tyr Ala Lys Pro Leu Tyr Lys Leu Lys Asn Thr			30
465	470	475	480
Val Ile Ile Ser Asn Pro Ala Val Ile Arg Leu His Gly Arg Asp Phe			35
485	490	495	
Leu Ile Ala His Gly Arg Gly Ile Glu Asp Val Val Ser Phe Val Pro			40
500	505	510	
Gly Leu Thr His His Lys Pro Gly Leu Pro Met Val Glu Leu Leu Lys			45
515	520	525	
Met Arg His Leu Ala Pro Thr Phe Gly Gly Lys Val Pro Ile Ala Pro			50
530	535	540	
Asp Pro Glu Asp Leu Leu Val Ile Glu Glu Val Pro Asp Leu Val Gln			55
545	550	555	560
Met Gly His Val His Val Tyr Asp Thr Ala Val Tyr Arg Gly Val Gln			60
565	570	575	
Leu Val Asn Ser Ala Thr Trp Gln Ala Gln Thr Glu Phe Gln Lys Met			65
580	585	590	
Val Asn Ile Val Pro Thr Pro Gly Leu Val Pro Ile Val Asp Val Glu			70
595	600	605	
Ser Ala Arg Val Ile Lys Val Leu Asp Phe Ser Arg Trp Cys			75
610	615	620	

## (2) ANGABEN ZU SEQ ID NO: 20:

## 5 (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 482 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

10

## (ii) ART DES MOLEKÜLS: Protein

15

## (vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Methanobacterium thermoautotrophicum

20

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

Met Asn Glu Ile Ile Gly Lys Phe Ala Arg Glu Gly Ile Leu Ile Glu

25

1 5 10 15

Asp Asn Ala Tyr Phe Arg Leu Arg Glu Met Asp Asp Pro Ala Ser Val

30

20 25 30

Ser Ser Glu Leu Ile Val Lys Ile Lys Ser Asn Gly Gly Lys Phe Thr

35 40 45

Val Leu Thr Ser Glu Met Leu Asp Glu Phe Phe Glu Ile Asp Asn Pro

50 55 60

Ala Glu Ile Lys Ala Arg Gly Pro Leu Met Val Pro Ala Glu Arg Asp

40

65 70 75 80

Phe Asp Phe Glu Val Ile Ser Asp Thr Ser Asn Arg Ser Tyr Thr Ser

45

85 90 95

Gly Glu Ile Gly Asp Met Ile Ala Tyr Phe Asn Ser Arg Tyr Ser Ser

50

100 105 110

Leu Lys Asn Leu Leu Ser Lys Arg Pro Glu Leu Lys Gly His Ile Pro

115 120 125

55

Ile Ala Asp Leu Arg Gly Gly Glu Asp Val Val Ser Ile Ile Gly Met

130 135 140

60

Val Asn Asp Val Arg Asn Thr Lys Asn Asn His Arg Ile Ile Glu Leu

145 150 155 160

DE 198 40 771 A 1

Glu	Asp	Asp	Thr	Gly	Glu	Ile	Ser	Val	Val	Val	His	Asn	Glu	Asn	His	
165																
Lys	Leu	Phe	Glu	Lys	Ser	Glu	Lys	Ile	Val	Arg	Asp	Glu	Val	Val	Gly	5
180																
Val	His	Gly	Thr	Lys	Lys	Gly	Arg	Phe	Val	Val	Ala	Ser	Glu	Ile	Phe	10
195																
His	Pro	Gly	Val	Pro	Arg	Ile	Gln	Glu	Lys	Glu	Met	Asp	Phe	Ser	Val	15
210																
Ala	Phe	Ile	Ser	Asp	Val	His	Ile	Gly	Ser	Gln	Thr	Phe	Leu	Glu	Asp	
225																
Ala	Phe	Met	Lys	Phe	Val	Lys	Trp	Ile	Asn	Gly	Asp	Phe	Gly	Ser	Glu	20
245																
Glu	Gln	Arg	Ser	Leu	Ala	Ala	Asp	Val	Lys	Tyr	Leu	Val	Val	Ala	Gly	25
260																
Asp	Ile	Val	Asp	Gly	Ile	Gly	Ile	Tyr	Pro	Gly	Gln	Glu	Lys	Glu	Leu	30
275																
Leu	Ile	Arg	Asp	Ile	His	Glu	Gln	Tyr	Glu	Glu	Ala	Ala	Arg	Leu	Phe	35
290																
Gly	Asp	Ile	Arg	Ser	Asp	Ile	Lys	Ile	Val	Met	Ile	Pro	Gly	Asn	His	
305																
Asp	Ser	Ser	Arg	Ile	Ala	Glu	Pro	Gln	Pro	Ala	Ile	Pro	Glu	Glu	Tyr	40
325																
Ala	Lys	Ser	Leu	Tyr	Ser	Ile	Arg	Asn	Ile	Glu	Phe	Leu	Ser	Asn	Pro	45
340																
Ser	Leu	Val	Ser	Leu	Asp	Gly	Val	Arg	Thr	Leu	Ile	Tyr	His	Gly	Arg	50
355																
Ser	Phe	Asp	Asp	Met	Ala	Met	Ser	Val	Asn	Gly	Leu	Ser	His	Glu	Arg	
370																
Ser	Asp	Leu	Ile	Met	Glu	Glu	Leu	Leu	Glu	Lys	Arg	His	Leu	Ala	Pro	55
385																
Ile	Tyr	Gly	Glu	Arg	Thr	Pro	Leu	Ala	Ser	Glu	Ile	Glu	Asp	His	Leu	60
405																

Val Ile Asp Glu Val Pro His Val Leu His Thr Gly His Val His Ile  
420 425 430

5 Asn Ala Tyr Lys Tyr Lys Gly Val His Leu Ile Asn Ser Gly Thr  
435 440 445

10 Phe Gln Ser Gln Thr Glu Phe Gln Lys Ile Tyr Asn Ile Val Pro Thr  
450 455 460

15 Cys Gly Gln Val Pro Val Leu Asn Arg Gly Val Met Lys Leu Leu Glu  
465 470 475 480

20 Phe Ser

20 (2) ANGABEN ZU SEQ ID NO: 21:

25 (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 613 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

30 (ii) ART DES MOLEKÜLS: Protein

35 (vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Pyrococcus furiosus

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

45 Met Asp Glu Phe Val Lys Ser Leu Leu Lys Ala Asn Tyr Leu Ile Thr  
1 5 10 15

50 Pro Ser Ala Tyr Tyr Leu Leu Arg Glu Tyr Tyr Glu Lys Gly Glu Phe  
20 25 30

55 Ser Ile Val Glu Leu Val Lys Phe Ala Arg Ser Arg Glu Ser Tyr Ile  
35 40 45

60 Ile Thr Asp Ala Leu Ala Thr Glu Phe Leu Lys Val Lys Gly Leu Glu  
50 55 60

65 Pro Ile Leu Pro Val Glu Thr Lys Gly Gly Phe Val Ser Thr Gly Glu  
65 70 75 80

## DE 198 40 771 A 1

Ser Gln Lys Glu Gln Ser Tyr Glu Glu Ser Phe Gly Thr Lys Glu Glu			
85	90	95	
Ile Ser Gln Glu Ile Lys Glu Gly Glu Ser Phe Ile Ser Thr Gly Ser			5
100	105	110	
Glu Pro Leu Glu Glu Glu Leu Asn Ser Ile Gly Ile Glu Glu Ile Gly			10
115	120	125	
Ala Asn Glu Glu Leu Val Ser Asn Gly Asn Asp Asn Gly Gly Glu Ala			
130	135	140	15
Ile Val Phe Asp Lys Tyr Gly Tyr Pro Met Val Tyr Ala Pro Glu Glu			
145	150	155	160
Ile Glu Val Glu Glu Lys Glu Tyr Ser Lys Tyr Glu Asp Leu Thr Ile			
165	170	175	
Pro Met Asn Pro Asp Phe Asn Tyr Val Glu Ile Lys Glu Asp Tyr Asp			25
180	185	190	
Val Val Phe Asp Val Arg Asn Val Lys Leu Lys Pro Pro Lys Val Lys			
195	200	205	30
Asn Gly Asn Gly Lys Glu Gly Glu Ile Ile Val Glu Ala Tyr Ala Ser			
210	215	220	35
Leu Phe Arg Ser Arg Leu Lys Lys Leu Arg Lys Ile Leu Arg Glu Asn			
225	230	235	240
Pro Glu Leu Asp Asn Val Val Asp Ile Gly Lys Leu Lys Tyr Val Lys			
245	250	255	40
Glu Asp Glu Thr Val Thr Ile Ile Gly Leu Val Asn Ser Lys Arg Glu			
260	265	270	45
Val Asn Lys Gly Leu Ile Phe Glu Ile Glu Asp Leu Thr Gly Lys Val			
275	280	285	50
Lys Val Phe Leu Pro Lys Asp Ser Glu Asp Tyr Arg Glu Ala Phe Lys			
290	295	300	55
Val Leu Pro Asp Ala Val Val Ala Phe Lys Gly Val Tyr Ser Lys Arg			
305	310	315	320
Gly Ile Leu Tyr Ala Asn Lys Phe Tyr Leu Pro Asp Val Pro Leu Tyr			
325	330	335	60
			65

## DE 198 40 771 A 1

Arg Arg Gln Lys Pro Pro Leu Glu Glu Lys Val Tyr Ala Ile Leu Ile  
 340 345 350

5 Ser Asp Ile His Val Gly Ser Lys Glu Phe Cys Glu Asn Ala Phe Ile  
 355 360 365

10 Lys Phe Leu Glu Trp Leu Asn Gly Asn Val Glu Thr Lys Glu Glu  
 370 375 380

15 Glu Ile Val Ser Arg Val Lys Tyr Leu Ile Ile Ala Gly Asp Val Val  
 385 390 395 400

20 Asp Gly Val Gly Val Tyr Pro Gly Gln Tyr Ala Asp Leu Thr Ile Pro  
 405 410 415

25 Asp Ile Phe Asp Gln Tyr Glu Ala Leu Ala Asn Leu Leu Ser His Val  
 420 425 430

30 Pro Lys His Ile Thr Met Phe Ile Ala Pro Gly Asn His Asp Ala Ala  
 435 440 445

35 Arg Gln Ala Ile Pro Gln Pro Glu Phe Tyr Lys Glu Tyr Ala Lys Pro  
 450 455 460

40 Ile Tyr Lys Leu Lys Asn Ala Val Ile Ile Ser Asn Pro Ala Val Ile  
 465 470 475 480

45 Arg Leu His Gly Arg Asp Phe Leu Ile Ala His Gly Arg Gly Ile Glu  
 485 490 495

50 Asp Val Val Gly Ser Val Pro Gly Leu Thr His His Lys Pro Gly Leu  
 500 505 510

55 Pro Met Val Glu Leu Leu Lys Met Arg His Val Ala Pro Met Phe Gly  
 515 520 525

60 Gly Lys Val Pro Ile Ala Pro Asp Pro Glu Asp Leu Leu Val Ile Glu  
 530 535 540

65 Glu Val Pro Asp Val Val His Met Gly His Val His Val Tyr Asp Ala  
 545 550 555 560

70 Val Val Tyr Arg Gly Val Gln Leu Val Asn Ser Ala Thr Trp Gln Ala  
 565 570 575

75 Gln Thr Glu Phe Gln Lys Met Val Asn Ile Val Pro Thr Pro Ala Lys  
 580 585 590

DE 198 40 771 A 1

Val Pro Val Val Asp Ile Asp Thr Ala Lys Val Val Lys Val Leu Asp  
 595 600 605

Phe Ser Gly Trp Cys  
 610

5

(2) ANGABEN ZU SEQ ID NO: 22:

10

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1107 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

15

(ii) ART DES MOLEKÜLS: Protein

20

(vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

30

Met Asp Gly Lys Arg Arg Pro Gly Pro Gly Pro Gly Val Pro Pro Lys  
 1 5 10 15

35

Arg Ala Arg Gly Gly Leu Trp Asp Asp Asp Asp Ala Pro Trp Pro Ser  
 20 25 30

40

Gln Phe Glu Glu Asp Leu Ala Leu Met Glu Glu Met Glu Ala Glu His  
 35 40 45

45

Arg Leu Gln Glu Gln Glu Glu Glu Leu Gln Ser Val Leu Glu Gly  
 50 55 60

45

Val Ala Asp Gly Gln Val Pro Pro Ser Ala Ile Asp Pro Arg Trp Leu  
 65 70 75 80

50

Arg Pro Thr Pro Pro Ala Leu Asp Pro Gln Thr Glu Pro Leu Ile Phe  
 85 90 95

55

Gln Gln Leu Glu Ile Asp His Tyr Val Gly Pro Ala Gln Pro Val Pro  
 100 105 110

55

Gly Gly Pro Pro Pro Ser Arg Gly Ser Val Pro Val Leu Arg Ala Phe  
 115 120 125

60

65

DE 198 40 771 A 1

Gly Val Thr Asp Glu Gly Phe Ser Val Cys Cys His Ile His Gly Phe  
130 135 140

5 Ala Pro Tyr Phe Tyr Thr Pro Ala Pro Pro Gly Phe Gly Pro Glu His  
145 150 155 160

10 Met Gly Asp Leu Gln Arg Glu Leu Asn Leu Ala Ile Ser Arg Asp Ser  
165 170 175

15 Arg Gly Gly Arg Glu Leu Thr Gly Pro Ala Val Leu Ala Val Glu Leu  
180 185 190

20 Cys Ser Arg Glu Ser Met Phe Gly Tyr His Gly His Gly Pro Ser Pro  
195 200 205

25 Phe Leu Arg Ile Thr Val Ala Leu Pro Arg Leu Val Ala Pro Ala Arg  
210 215 220

30 Arg Leu Leu Glu Gln Gly Ile Arg Val Ala Gly Leu Gly Thr Pro Ser  
225 230 235 240

35 Phe Ala Pro Tyr Glu Ala Asn Val Asp Phe Glu Ile Arg Phe Met Val  
245 250 255

40 Asp Thr Asp Ile Val Gly Cys Asn Trp Leu Glu Leu Pro Ala Gly Lys  
260 265 270

45 Tyr Ala Leu Arg Leu Lys Glu Lys Ala Thr Gln Cys Gln Leu Glu Ala  
275 280 285

50 Asp Val Leu Trp Ser Asp Val Val Ser His Pro Pro Glu Gly Pro Trp  
290 295 300

55 Gln Arg Ile Ala Pro Leu Arg Val Leu Ser Phe Asp Ile Glu Cys Ala  
305 310 315 320

60 Gly Arg Lys Gly Ile Phe Pro Glu Pro Glu Arg Asp Pro Val Ile Gln  
325 330 335

Ile Cys Ser Leu Gly Leu Arg Trp Gly Glu Pro Glu Pro Phe Leu Arg  
340 345 350

65 Leu Ala Leu Thr Leu Arg Pro Cys Ala Pro Ile Leu Gly Ala Lys Val  
355 360 365

70 Gln Ser Tyr Glu Lys Glu Glu Asp Leu Leu Gln Ala Trp Ser Thr Phe  
370 375 380

DE 198 40 771 A 1

Ile Arg Ile Met Asp Pro Asp Val Ile Thr Gly Asn Ile Gln Asn			
385	390	395	400
Phe Asp Leu Pro Tyr Leu Ile Ser Arg Ala Gln Thr Leu Lys Val Gln			
405	410	415	
Thr Phe Pro Phe Leu Gly Arg Val Ala Gly Leu Cys Ser Asn Ile Arg			
420	425	430	
Asp Ser Ser Phe Gln Ser Lys Gln Thr Gly Arg Arg Asp Thr Lys Val			
435	440	445	
Val Ser Met Val Gly Arg Val Gln Met Asp Met Leu Gln Val Leu Leu			
450	455	460	
Arg Glu Tyr Lys Leu Arg Ser His Thr Leu Asn Ala Val Ser Phe His			
465	470	475	480
Phe Leu Gly Glu Gln Lys Glu Asp Val Gln His Ser Ile Ile Thr Asp			
485	490	495	
Leu Gln Asn Gly Asn Asp Gln Thr Arg Arg Arg Leu Ala Val Tyr Cys			
500	505	510	
Leu Lys Asp Ala Tyr Leu Pro Leu Arg Leu Leu Glu Arg Leu Met Val			
515	520	525	
Leu Val Asn Ala Val Glu Met Ala Arg Val Thr Gly Val Pro Leu Ser			
530	535	540	
Tyr Leu Leu Ser Arg Gly Gln Gln Val Lys Val Val Ser Gln Leu Leu			
545	550	555	560
Arg Gln Ala Met His Glu Gly Leu Leu Met Pro Val Val Lys Ser Glu			
565	570	575	
Gly Gly Glu Asp Tyr Thr Gly Ala Thr Val Ile Glu Pro Leu Lys Gly			
580	585	590	
Tyr Tyr Asp Val Pro Ile Ala Thr Leu Asp Phe Ser Ser Leu Tyr Pro			
595	600	605	
Ser Ile Met Met Ala His Asn Leu Cys Tyr Thr Thr Leu Leu Arg Pro			
610	615	620	
Gly Thr Ala Gln Lys Leu Gly Leu Thr Glu Asp Gln Phe Ile Arg Thr			
625	630	635	640

DE 198 40 771 A 1

Pro Thr Gly Asp Glu Phe Val Lys Thr Ser Val Arg Lys Gly Leu Leu  
645 650 655

<sup>5</sup> Pro Gln Ile Leu Glu Asn Leu Leu Ser Ala Arg Lys Arg Ala Lys Ala  
660 665 670

<sup>10</sup> Glu Leu Ala Lys Glu Thr Asp Pro Leu Arg Arg Gln Val Leu Asp Gly  
675 680 685

<sup>15</sup> Arg Gln Leu Ala Leu Lys Val Ser Ala Asn Ser Val Tyr Gly Phe Thr  
690 695 700

<sup>20</sup> Gly Ala Gln Val Gly Lys Leu Pro Cys Leu Glu Ile Ser Gln Ser Val  
705 710 715 720

Thr Gly Phe Gly Arg Gln Met Ile Glu Lys Thr Lys Gln Leu Val Glu  
725 730 735

<sup>25</sup> Ser Lys Tyr Thr Val Glu Asn Gly Tyr Ser Thr Ser Ala Lys Val Val  
740 745 750

<sup>30</sup> Tyr Gly Asp Thr Asp Ser Val Met Cys Arg Phe Gly Val Ser Ser Val  
755 760 765

<sup>35</sup> Ala Glu Ala Met Ala Leu Gly Arg Glu Ala Ala Asp Trp Val Ser Gly  
770 775 780

His Phe Pro Ser Pro Ile Arg Leu Glu Phe Glu Lys Val Tyr Phe Pro  
785 790 795 800

<sup>40</sup> Tyr Leu Leu Ile Ser Lys Lys Arg Tyr Ala Gly Leu Leu Phe Ser Ser  
805 810 815

<sup>45</sup> Arg Pro Asp Ala His Asp Arg Met Asp Cys Lys Gly Leu Glu Ala Val  
820 825 830

<sup>50</sup> Arg Arg Asp Asn Cys Pro Leu Val Ala Asn Leu Val Thr Ala Ser Leu  
835 840 845

<sup>55</sup> Arg Arg Leu Leu Ile Asp Arg Asp Pro Glu Gly Ala Val Ala His Ala  
850 855 860

Gln Asp Val Ile Ser Asp Leu Leu Cys Asn Arg Ile Asp Ile Ser Gln  
865 870 875 880

<sup>60</sup> Leu Val Ile Thr Lys Glu Leu Thr Arg Ala Ala Ser Asp Tyr Ala Gly  
885 890 895

DE 198 40 771 A 1

Lys	Gln	Ala	His	Val	Glu	Leu	Ala	Glu	Arg	Met	Arg	Lys	Arg	Asp	Pro		
900																	
Gly	S	r	Ala	Pro	Ser	Leu	Gly	Asp	Arg	Val	Pro	Tyr	Val	Ile	Ile	Ser	5
915																	
Ala	Ala	Lys	Gly	Val	Ala	Ala	Tyr	Met	Lys	Ser	Glu	Asp	Pro	Leu	Phe		10
930																	
Val	Leu	Glu	His	Ser	Leu	Pro	Ile	Asp	Thr	Gln	Tyr	Tyr	Leu	Glu	Gln		15
945																	
Gln	Leu	Ala	Lys	Pro	Leu	Leu	Arg	Ile	Phe	Glu	Pro	Ile	Leu	Gly	Glu		20
965																	
Gly	Arg	Ala	Glu	Ala	Val	Leu	Leu	Arg	Gly	Asp	His	Thr	Arg	Cys	Lys		25
980																	
Thr	Val	Leu	Thr	Gly	Lys	Val	Gly	Gly	Leu	Leu	Ala	Phe	Ala	Lys	Arg		30
995																	
Arg	Asn	Cys	Cys	Ile	Gly	Cys	Arg	Thr	Val	Leu	Ser	His	Gln	Gly	Ala		35
1010																	
Val	Cys	Glu	Phe	Cys	Gln	Pro	Arg	Glu	Ser	Glu	Leu	Tyr	Gln	Lys	Glu		40
1025																	
Val	Ser	His	Leu	Asn	Ala	Leu	Glu	Glu	Arg	Phe	Ser	Arg	Leu	Trp	Thr		45
1045																	
Gln	Cys	Gln	Arg	Cys	Gln	Gly	Ser	Leu	His	Glu	Asp	Val	Ile	Cys	Thr		50
1060																	
Ser	Arg	Asp	Cys	Pro	Ile	Phe	Tyr	Met	Arg	Lys	Lys	Val	Arg	Lys	Asp		55
1075																	
Leu	Glu	Asp	Gln	Glu	Gln	Leu	Leu	Arg	Arg	Phe	Gly	Pro	Pro	Gly	Pro		60
1090																	
Glu	Ala	Trp															65
1105																	

(2) ANGABEN ZU SEQ ID NO: 23:

- (i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 781 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang

65

(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Archaeoglobus fulgidus

10

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

15

Met Glu Arg Val Glu Gly Trp Leu Ile Asp Ala Asp Tyr Glu Thr Ile  
1 5 10 15

20

Gly Gly Lys Ala Val Val Arg Leu Trp Cys Lys Asp Asp Gln Gly Ile  
20 25 30

25

Phe Val Ala Tyr Asp Tyr Asn Phe Asp Pro Tyr Phe Tyr Val Ile Gly  
35 40 45

30

Val Asp Glu Asp Ile Leu Lys Asn Ala Ala Thr Ser Thr Arg Arg Glu  
50 55 60

35

Val Ile Lys Leu Lys Ser Phe Glu Lys Ala Gln Leu Lys Thr Leu Gly  
65 70 75 80

40

Arg Glu Val Glu Gly Tyr Ile Val Tyr Ala His His Pro Gln His Val  
85 90 95

45

Pro Lys Leu Arg Asp Tyr Leu Ser Gln Phe Gly Asp Val Arg Glu Ala  
100 105 110

50

Asp Ile Pro Phe Ala Tyr Arg Tyr Leu Ile Asp Lys Asp Leu Ala Cys  
115 120 125

55

Met Asp Gly Ile Ala Ile Glu Gly Glu Lys Gln Gly Gly Val Ile Arg  
130 135 140

60

Ser Tyr Lys Ile Glu Lys Val Glu Arg Ile Pro Arg Met Glu Phe Pro  
145 150 155 160

65

Glu Leu Lys Met Leu Val Phe Asp Cys Glu Met Leu Ser Ser Phe Gly  
165 170 175

70

Met Pro Glu Pro Glu Lys Asp Pro Ile Ile Val Ile Ser Val Lys Thr  
180 185 190

65

DE 198 40 771 A 1

Asn	Asp	Asp	Asp	Glu	Ile	Ile	Leu	Thr	Gly	Asp	Glu	Arg	Lys	Ile	Ile	
195				200												5
Ser	Asp	Phe	Val	Lys	Leu	Ile	Lys	Ser	Tyr	Asp	Pro	Asp	Ile	Ile	Val	
210		215			220											
Gly	Tyr	Asn	Gln	Asp	Ala	Phe	Asp	Trp	Pro	Tyr	Leu	Arg	Lys	Arg	Ala	
225		230		235				240								10
Glu	Arg	Trp	Asn	Ile	Pro	Leu	Asp	Val	Gly	Arg	Asp	Gly	Ser	Asn	Val	
245		250			255											15
Val	Phe	Arg	Gly	Gly	Arg	Pro	Lys	Ile	Thr	Gly	Arg	Leu	Asn	Val	Asp	
260		265			270											20
Leu	Tyr	Asp	Ile	Ala	Met	Arg	Ile	Ser	Asp	Ile	Lys	Ile	Lys	Leu		
275		280			285											
Glu	Asn	Val	Ala	Glu	Phe	Leu	Gly	Thr	Lys	Ile	Glu	Ile	Ala	Asp	Ile	
290		295			300											25
Glu	Ala	Lys	Asp	Ile	Tyr	Arg	Tyr	Trp	Ser	Arg	Gly	Glu	Lys			
305		310		315			320									30
Val	Leu	Asn	Tyr	Ala	Arg	Gln	Asp	Ala	Ile	Asn	Thr	Tyr	Leu	Ile	Ala	
325		330			335											35
Lys	Glu	Leu	Leu	Pro	Met	His	Tyr	Glu	Leu	Ser	Lys	Met	Ile	Arg	Leu	
340		345			350											
Pro	Val	Asp	Asp	Val	Thr	Arg	Met	Gly	Arg	Gly	Lys	Gln	Val	Asp	Trp	
355		360			365											40
Leu	Leu	Leu	Ser	Glu	Ala	Lys	Ile	Gly	Glu	Ile	Ala	Pro	Asn	Pro		
370		375			380											45
Pro	Glu	His	Ala	Glu	Ser	Tyr	Glu	Gly	Ala	Phe	Val	Leu	Glu	Pro	Glu	
385		390			395			400								50
Arg	Gly	Leu	His	Glu	Asn	Val	Ala	Cys	Leu	Asp	Phe	Ala	Ser	Met	Tyr	
405		410			415											55
Pro	Ser	Ile	Met	Ile	Ala	Phe	Asn	Ile	Ser	Pro	Asp	Thr	Tyr	Gly	Cys	
420		425			430											
Arg	Asp	Asp	Cys	Tyr	Glu	Ala	Pro	Glu	Val	Gly	His	Lys	Phe	Arg	Lys	
435		440			445											60
																65

DE 198 40 771 A 1

Ser Pro Asp Gly Phe Phe Lys Arg Ile Leu Arg Met Leu Ile Glu Lys  
450 455 460

5 Arg Arg Glu Leu Lys Val Glu Leu Lys Asn Leu Ser Pro Glu Ser Ser  
465 470 475 480

10 Glu Tyr Lys Leu Leu Asp Ile Lys Gln Gln Thr Leu Lys Val Leu Thr  
485 490 495

15 Asn Ser Phe Tyr Gly Tyr Met Gly Trp Asn Leu Ala Arg Trp Tyr Cys  
500 505 510

20 His Pro Cys Ala Glu Ala Thr Thr Ala Trp Gly Arg His Phe Ile Arg  
515 520 525

25 Thr Ser Ala Lys Ile Ala Glu Ser Met Gly Phe Lys Val Leu Tyr Gly  
530 535 540

30 25 Asp Thr Asp Ser Ile Phe Val Thr Lys Ala Gly Met Thr Lys Glu Asp  
545 550 555 560

35 Val Asp Arg Leu Ile Asp Lys Leu His Glu Glu Leu Pro Ile Gln Ile  
565 570 575

40 Glu Val Asp Glu Tyr Tyr Ser Ala Ile Phe Phe Val Glu Lys Lys Arg  
580 585 590

45 Tyr Ala Gly Leu Thr Glu Asp Gly Arg Leu Val Val Lys Gly Leu Glu  
595 600 605

50 Val Arg Arg Gly Asp Trp Cys Glu Leu Ala Lys Lys Val Gln Arg Glu  
610 615 620

55 45 Val Ile Glu Val Ile Leu Lys Glu Lys Asn Pro Glu Lys Ala Leu Ser  
625 630 635 640

60 Leu Val Lys Asp Val Ile Leu Arg Ile Lys Glu Gly Lys Val Ser Leu  
645 650 655

65 Glu Glu Val Val Ile Tyr Lys Gly Leu Thr Lys Lys Pro Ser Lys Tyr  
660 665 670

70 Glu Ser Met Gln Ala His Val Lys Ala Ala Leu Lys Ala Arg Glu Met  
675 680 685

75 Gly Ile Ile Tyr Pro Val Ser Ser Lys Ile Gly Tyr Val Ile Val Lys  
690 695 700

DE 198 40 771 A 1

Gly Ser Gly Asn Ile Gly Asp Arg Ala Tyr Pro Ile Asp Leu Ile Glu				
705	710	715	720	5
Asp Phe Asp Gly Glu Asn Leu Arg Ile Lys Thr Lys Ser Gly Ile Glu				
725	730	735		
Ile Lys Lys Leu Asp Lys Asp Tyr Tyr Ile Asp Asn Gln Ile Ile Pro				10
740	745	750		
Ser Val Leu Arg Ile Leu Glu Arg Phe Gly Tyr Thr Glu Ala Ser Leu				15
755	760	765		
Lys Gly Ser Ser Gln Met Ser Leu Asp Ser Phe Phe Ser				
770	775	780		20
(2) ANGABEN ZU SEQ ID NO: 24:				
(i) SEQUENZKENNZEICHEN:				
(A) LÄNGE: 1634 Aminosäuren				
(B) ART: Aminosäure				
(C) STRANGFORM: Einzelstrang				
(D) TOPOLOGIE: linear				
(ii) ART DES MOLEKÜLS: Protein				
(vi) URSPRÜNLICHE HERKUNFT:				
(A) ORGANISMUS: Methanococcus jannaschii				
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:				
Met Gly Met Ser Met Gly Lys Ile Lys Ile Asp Ala Leu Ile Asp Asn				40
1	5	10	15	45
Thr Tyr Lys Thr Ile Glu Asp Lys Ala Val Ile Tyr Leu Tyr Leu Ile				
20	25	30		50
Asn Ser Ile Leu Lys Asp Arg Asp Phe Lys Pro Tyr Phe Tyr Val Glu				
35	40	45		
Leu His Lys Glu Lys Val Glu Asn Glu Asp Ile Glu Lys Ile Lys Glu				55
50	55	60		
Phe Leu Leu Lys Asn Asp Leu Leu Lys Phe Val Glu Asn Ile Glu Val				
65	70	75	80	60

65

## DE 198 40 771 A 1

Val Lys Lys Ile Ile Leu Arg Lys Glu Lys Glu Val Ile Lys Ile Ile		
85	90	95
5 Ala Thr His Pro Gln Lys Val Pro Lys Leu Arg Lys Ile Lys Glu Cys		
100	105	110
10 Glu Ile Val Lys Glu Ile Tyr Glu His Asp Ile Pro Phe Ala Lys Arg		
115	120	125
15 Tyr Leu Ile Asp Asn Glu Ile Ile Pro Met Thr Tyr Trp Asp Phe Glu		
130	135	140
20 Asn Lys Lys Pro Val Ser Ile Glu Ile Pro Lys Leu Lys Ser Val Ala		
145	150	155
25 Phe Asp Met Glu Val Tyr Asn Arg Asp Thr Glu Pro Asn Pro Glu Arg		
165	170	175
30 Asp Pro Ile Leu Met Ala Ser Phe Trp Asp Glu Asn Gly Gly Lys Val		
180	185	190
35 Ile Thr Tyr Lys Glu Phe Asn His Pro Asn Ile Glu Val Val Lys Asn		
195	200	205
40 Glu Lys Glu Leu Ile Lys Lys Ile Ile Glu Thr Leu Lys Glu Tyr Asp		
210	215	220
45 Val Ile Tyr Thr Tyr Asn Gly Asp Asn Phe Asp Phe Pro Tyr Leu Lys		
225	230	235
50 Ala Arg Ala Lys Ile Tyr Gly Ile Asp Ile Asn Leu Gly Lys Asp Gly		
245	250	255
55 Glu Glu Leu Lys Ile Lys Arg Gly Gly Met Glu Tyr Arg Ser Tyr Ile		
260	265	270
60 Pro Gly Arg Val His Ile Asp Leu Tyr Pro Ile Ser Arg Arg Leu Leu		
275	280	285
65 Lys Leu Thr Lys Tyr Thr Leu Glu Asp Val Val Tyr Asn Leu Phe Gly		
290	295	300
70 Ile Glu Lys Leu Lys Ile Pro His Thr Lys Ile Val Asp Tyr Trp Ala		
305	310	315
75 Asn Asn Asp Lys Thr Leu Ile Glu Tyr Ser Leu Gln Asp Ala Lys Tyr		
325	330	335

## DE 198 40 771 A 1

Thr Tyr Lys Ile Gly Lys Tyr Phe Phe Pro Leu Glu Val Met Phe Ser				
340	345	350		
Arg Ile Val Asn Gln Thr Pro Phe Glu Ile Thr Arg Met Ser Ser Gly				5
355	360	365		
Gln Met Val Glu Tyr Leu Leu Met Lys Arg Ala Phe Lys Glu Asn Met				10
370	375	380		
Ile Val Pro Asn Lys Pro Asp Glu Glu Glu Tyr Arg Arg Arg Val Leu				15
385	390	395	400	
Thr Thr Tyr Glu Gly Gly Tyr Val Lys Glu Pro Glu Lys Gly Met Phe				20
405	410	415		
Glu Asp Ile Ile Ser Met Asp Phe Arg Cys His Pro Lys Gly Thr Lys				
420	425	430		
Val Val Val Lys Gly Lys Gly Ile Val Asn Ile Glu Asp Val Lys Glu				25
435	440	445		
Gly Asn Tyr Val Leu Gly Ile Asp Gly Trp Gln Lys Val Lys Lys Val				30
450	455	460		
Trp Lys Tyr Glu Tyr Glu Gly Glu Leu Ile Asn Val Asn Gly Leu Lys				35
465	470	475	480	
Cys Thr Pro Asn His Lys Ile Pro Leu Arg Tyr Lys Ile Lys His Lys				
485	490	495		
Lys Ile Asn Lys Asn Asp Tyr Leu Val Arg Asp Ile Tyr Ala Lys Ser				40
500	505	510		
Leu Leu Thr Lys Phe Lys Gly Glu Lys Leu Ile Leu Cys Lys Asp				45
515	520	525		
Phe Glu Thr Ile Gly Asn Tyr Glu Lys Tyr Ile Asn Asp Met Asp Glu				
530	535	540		
Asp Phe Ile Leu Lys Ser Glu Leu Ile Gly Ile Leu Leu Ala Glu Gly				
545	550	555	560	
His Leu Leu Arg Arg Asp Ile Glu Tyr Phe Asp Ser Ser Arg Gly Lys				55
565	570	575		
Lys Arg Ile Ser His Gln Tyr Arg Val Glu Ile Thr Val Asn Glu Asp				60
580	585	590		

65

DE 198 40 771 A 1

Glu Lys Asp Phe Ile Glu Lys Ile Lys Tyr Ile Phe Lys Lys Leu Phe  
595 600 605

5 Asn Tyr Glu Leu Tyr Val Arg Arg Lys Lys Gly Thr Lys Ala Ile Thr  
610 615 620

10 Leu Gly Cys Ala Lys Lys Asp Ile Tyr Leu Lys Ile Glu Glu Ile Leu  
625 630 635 640

15 Lys Asn Lys Glu Lys Tyr Leu Pro Asn Ala Ile Leu Arg Gly Phe Phe  
645 650 655

20 Glu Gly Asp Gly Tyr Val Asn Thr Val Arg Arg Ala Val Val Val Asn  
660 665 670

25 Gln Gly Thr Asn Asn Tyr Asp Lys Ile Lys Phe Ile Ala Ser Leu Leu  
675 680 685

30 Asp Arg Leu Gly Ile Lys Tyr Ser Phe Tyr Thr Tyr Ser Tyr Glu Glu  
690 695 700

35 Arg Gly Lys Lys Leu Lys Arg Tyr Val Ile Glu Ile Phe Ser Lys Gly  
705 710 715 720

40 Asp Leu Ile Lys Phe Ser Ile Leu Ile Ser Phe Ile Ser Arg Arg Lys  
725 730 735

45 Asn Asn Leu Leu Asn Glu Ile Ile Arg Gln Lys Thr Leu Tyr Lys Ile  
740 745 750

50 Gly Asp Tyr Gly Phe Tyr Asp Leu Asp Asp Val Cys Val Ser Leu Glu  
755 760 765

55 Ser Tyr Lys Gly Glu Val Tyr Asp Leu Thr Leu Glu Gly Arg Pro Tyr  
770 775 780

60 Tyr Phe Ala Asn Gly Ile Leu Thr His Asn Ser Leu Tyr Pro Ser Ile  
785 790 795 800

65 Ile Ile Ser Tyr Asn Ile Ser Pro Asp Thr Leu Asp Cys Glu Cys Cys  
805 810 815

70 Lys Asp Val Ser Glu Lys Ile Leu Gly His Trp Phe Cys Lys Lys Lys  
820 825 830

75 Glu Gly Leu Ile Pro Lys Thr Leu Arg Asn Leu Ile Glu Arg Arg Ile  
835 840 845

DE 198 40 771 A 1

Asn Ile Lys Arg Arg Met Lys Lys Met Ala Glu Ile Gly Glu Ile Asn			
850	855	860	5
Glu Glu Tyr Asn Leu Leu Asp Tyr Glu Gln Lys Ser Leu Lys Ile Leu			
865	870	875	880
Ala Asn Ser Ile Leu Pro Asp Glu Tyr Leu Thr Ile Ile Glu Glu Asp			10
885	890	895	
Gly Ile Lys Val Val Lys Ile Gly Glu Tyr Ile Asp Asp Leu Met Arg			15
900	905	910	
Lys His Lys Asp Lys Ile Lys Phe Ser Gly Ile Ser Glu Ile Leu Glu			
915	920	925	20
Thr Lys Asn Leu Lys Thr Phe Ser Phe Asp Lys Ile Thr Lys Lys Cys			
930	935	940	
Glu Ile Lys Lys Val Lys Ala Leu Ile Arg His Pro Tyr Phe Gly Lys			25
945	950	955	960
Ala Tyr Lys Ile Lys Leu Arg Ser Gly Arg Thr Ile Lys Val Thr Arg			30
965	970	975	
Gly His Ser Leu Phe Lys Tyr Glu Asn Gly Lys Ile Val Glu Val Lys			35
980	985	990	
Gly Asp Asp Val Arg Phe Gly Asp Leu Ile Val Val Pro Lys Lys Leu			
995	1000	1005	40
Thr Cys Val Asp Lys Glu Val Val Ile Asn Ile Pro Lys Arg Leu Ile			
1010	1015	1020	
Asn Ala Asp Glu Glu Glu Ile Lys Asp Leu Val Ile Thr Lys His Lys			45
1025	1030	1035	1040
Asp Lys Ala Phe Phe Val Lys Leu Lys Lys Thr Leu Glu Asp Ile Glu			50
1045	1050	1055	
Asn Asn Lys Leu Lys Val Ile Phe Asp Asp Cys Ile Leu Tyr Leu Lys			
1060	1065	1070	55
Glu Leu Gly Leu Ile Asp Tyr Asn Ile Ile Lys Lys Ile Asn Lys Val			
1075	1080	1085	60
Asp Ile Lys Ile Leu Asp Glu Glu Lys Phe Lys Ala Tyr Lys Lys Tyr			
1090	1095	1100	

DE 198 40 771 A 1

Phe Asp Thr Val Ile Glu His Gly Asn Phe Lys Lys Gly Arg Cys Asn  
 1105 1110 1115 1120  
 5 Ile Gln Tyr Ile Lys Ile Lys Asp Tyr Ile Ala Asn Ile Pro Asp Lys  
 1125 1130 1135  
 10 Glu Phe Glu Asp Cys Glu Ile Gly Ala Tyr Ser Gly Lys Ile Asn Ala  
 1140 1145 1150  
 15 Leu Leu Lys Leu Asp Glu Lys Leu Ala Lys Phe Leu Gly Phe Phe Val  
 1155 1160 1165  
 20 Thr Arg Gly Arg Leu Lys Lys Gln Lys Leu Lys Gly Glu Thr Val Tyr  
 1170 1175 1180  
 25 Glu Ile Ser Val Tyr Lys Ser Leu Pro Glu Tyr Gln Lys Glu Ile Ala  
 1185 1190 1195 1200  
 30 Glu Thr Phe Lys Glu Val Phe Gly Ala Gly Ser Met Val Lys Asp Lys  
 1205 1210 1215  
 35 Val Thr Met Asp Asn Lys Ile Val Tyr Leu Val Leu Lys Tyr Ile Phe  
 1220 1225 1230  
 Lys Cys Gly Asp Lys Asp Lys Lys His Ile Pro Glu Glu Leu Phe Leu  
 1235 1240 1245  
 40 Ala Ser Glu Ser Val Ile Lys Ser Phe Leu Asp Gly Phe Leu Lys Ala  
 1250 1255 1260  
 Lys Lys Asn Ser His Lys Gly Thr Ser Thr Phe Met Ala Lys Asp Glu  
 1265 1270 1275 1280  
 45 Lys Tyr Leu Asn Gln Leu Met Ile Leu Phe Asn Leu Val Gly Ile Pro  
 1285 1290 1295  
 50 Thr Arg Phe Thr Pro Val Lys Asn Lys Gly Tyr Lys Leu Thr Leu Asn  
 1300 1305 1310  
 Pro Lys Tyr Gly Thr Val Lys Asp Leu Met Leu Asp Glu Val Lys Glu  
 55 1315 1320 1325  
 Ile Glu Ala Phe Glu Tyr Ser Gly Tyr Val Tyr Asp Leu Ser Val Glu  
 1330 1335 1340  
 60 Asp Asn Glu Asn Phe Leu Val Asn Asn Ile Tyr Ala His Asn Ser Val  
 1345 1350 1355 1360

## DE 198 40 771 A 1

Tyr Gly Tyr Leu Ala Phe Pro Arg Ala Arg Phe Tyr Ser Arg Glu Cys			
1365	1370	1375	
Ala Glu Ile Val Thr Tyr Leu Gly Arg Lys Tyr Ile Leu Glu Thr Val			5
1380	1385	1390	
Lys Glu Ala Glu Lys Phe Gly Phe Lys Val Leu Tyr Ile Asp Thr Asp			10
1395	1400	1405	
Gly Phe Tyr Ala Ile Trp Lys Glu Lys Ile Ser Lys Glu Glu Leu Ile			
1410	1415	1420	15
Lys Lys Ala Met Glu Phe Val Glu Tyr Ile Asn Ser Lys Leu Pro Gly			
1425	1430	1435	1440
Thr Met Glu Leu Glu Phe Glu Gly Tyr Phe Lys Arg Gly Ile Phe Val			20
1445	1450	1455	
Thr Lys Lys Arg Tyr Ala Leu Ile Asp Glu Asn Gly Arg Val Thr Val			25
1460	1465	1470	
Lys Gly Leu Glu Phe Val Arg Arg Asp Trp Ser Asn Ile Ala Lys Ile			30
1475	1480	1485	
Thr Gln Arg Arg Val Leu Glu Ala Leu Leu Val Glu Gly Ser Ile Glu			
1490	1495	1500	35
Lys Ala Lys Lys Ile Ile Gln Asp Val Ile Lys Asp Leu Arg Glu Lys			
1505	1510	1515	1520
Lys Ile Lys Lys Glu Asp Leu Ile Ile Tyr Thr Gln Leu Thr Lys Asp			40
1525	1530	1535	
Pro Lys Glu Tyr Lys Thr Thr Ala Pro His Val Glu Ile Ala Lys Lys			45
1540	1545	1550	
Leu Met Arg Glu Gly Lys Arg Ile Lys Val Gly Asp Ile Ile Gly Tyr			
1555	1560	1565	50
Ile Ile Val Lys Gly Thr Lys Ser Ile Ser Glu Arg Ala Lys Leu Pro			
1570	1575	1580	
Glu Glu Val Asp Ile Asp Asp Ile Asp Val Asn Tyr Tyr Ile Asp Asn			
1585	1590	1595	1600
Gln Ile Leu Pro Pro Val Leu Arg Ile Met Glu Ala Val Gly Val Ser			60
1605	1610	1615	

65

DE 198 40 771 A 1

Lys Asn Glu Leu Lys Lys Glu Gly Ala Gln Leu Thr Leu Asp Lys Phe  
1620 1625 1630

5 Phe Lys

10 (2) ANGABEN ZU SEQ ID NO: 25:

15 (i) SEQUENZKENNZEICHEN:  
(A) LÄNGE: 1235 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

20 (ii) ART DES MOLEKÜLS: Protein

25 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Pyrococcus horikoshii

30 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

35 Met Ile Leu Asp Ala Asp Tyr Ile Thr Glu Asp Gly Lys Pro Ile Ile  
1 5 10 15

Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Val Glu Tyr Asp Arg  
20 25 30

40 Asn Phe Arg Pro Tyr Ile Tyr Ala Leu Leu Arg Asp Asp Ser Ala Ile  
35 40 45

45 Asp Glu Ile Lys Lys Ile Thr Ala Gln Arg His Gly Lys Val Val Arg  
50 55 60

50 Ile Val Glu Thr Glu Lys Ile Gln Arg Lys Phe Leu Gly Arg Pro Ile  
65 70 75 80

65 Glu Val Trp Lys Leu Tyr Leu Glu His Pro Gln Asp Val Pro Ala Ile  
85 90 95

55 Arg Asp Lys Ile Arg Glu His Pro Ala Val Val Asp Ile Phe Glu Tyr  
100 105 110

60 Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Thr Pro  
115 120 125

DE 198 40 771 A 1

DE 198 40 771 A 1

Tyr Glu Gly Gly Tyr Val Lys Glu Pro Glu Lys Gly Leu Trp Glu Gly  
 385 390 395 400

5 Ile Val Ser Leu Asp Phe Arg Ser Leu Tyr Pro Ser Ile Ile Ile Thr  
 405 410 415

10 His Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Glu Glu Tyr  
 420 425 430

15 Asp Val Ala Pro Lys Val Gly His Arg Phe Cys Lys Asp Phe Pro Gly  
 435 440 445

20 Phe Ile Pro Ser Leu Leu Gly Gln Leu Leu Glu Glu Arg Gln Lys Ile  
 450 455 460

25 Lys Lys Arg Met Lys Glu Ser Lys Asp Pro Val Glu Lys Lys Leu Leu  
 465 470 475 480

30 Asp Tyr Arg Gln Arg Ala Ile Lys Ile Leu Ala Asn Ser Ile Leu Pro  
 485 490 495

35 Asp Glu Trp Leu Pro Ile Val Glu Asn Glu Lys Val Arg Phe Val Lys  
 500 505 510

40 Ile Gly Asp Phe Ile Asp Arg Glu Ile Glu Glu Asn Ala Glu Arg Val  
 515 520 525

45 Lys Arg Asp Gly Glu Thr Glu Ile Leu Glu Val Lys Asp Leu Lys Ala  
 530 535 540

50 Leu Ser Phe Asn Arg Glu Thr Lys Lys Ser Glu Leu Lys Lys Val Lys  
 545 550 555 560

55 Ala Leu Ile Arg His Arg Tyr Ser Gly Lys Val Tyr Ser Ile Lys Leu  
 565 570 575

60 Lys Ser Gly Arg Arg Ile Lys Ile Thr Ser Gly His Ser Leu Phe Ser  
 580 585 590

65 Val Lys Asn Gly Lys Leu Val Lys Val Arg Gly Asp Glu Leu Lys Pro  
 595 600 605

70 Gly Asp Leu Val Val Val Pro Gly Arg Leu Lys Leu Pro Glu Ser Lys  
 610 615 620

75 Gln Val Leu Asn Leu Val Glu Leu Leu Leu Lys Leu Pro Glu Glu Glu  
 625 630 635 640

DE 198 40 771 A 1

Thr Ser Asn Ile Val Met Met Ile Pro Val Lys Gly Arg Lys Asn Phe				
645	650	655		
Phe Lys Gly Met Leu Lys Thr Leu Tyr Trp Ile Phe Gly Glu Gly Glu				5
660	665	670		
Arg Pro Arg Thr Ala Gly Arg Tyr Leu Lys His Leu Glu Arg Leu Gly				10
675	680	685		
Tyr Val Lys Leu Lys Arg Arg Gly Cys Glu Val Leu Asp Trp Glu Ser				15
690	695	700		
Leu Lys Arg Tyr Arg Lys Leu Tyr Glu Thr Leu Ile Lys Asn Leu Lys				
705	710	715	720	
Tyr Asn Gly Asn Ser Arg Ala Tyr Met Val Glu Phe Asn Ser Leu Arg				20
725	730	735		
Asp Val Val Ser Leu Met Pro Ile Glu Glu Leu Lys Glu Trp Ile Ile				25
740	745	750		
Gly Glu Pro Arg Gly Pro Lys Ile Gly Thr Phe Ile Asp Val Asp Asp				30
755	760	765		
Ser Phe Ala Lys Leu Leu Gly Tyr Tyr Ile Ser Ser Gly Asp Val Glu				35
770	775	780		
Lys Asp Arg Val Lys Phe His Ser Lys Asp Gln Asn Val Leu Glu Asp				40
785	790	795	800	
Ile Ala Lys Leu Ala Glu Lys Leu Phe Gly Lys Val Arg Arg Gly Arg				
805	810	815		
Gly Tyr Ile Glu Val Ser Gly Lys Ile Ser His Ala Ile Phe Arg Val				45
820	825	830		
Leu Ala Glu Gly Lys Arg Ile Pro Glu Phe Ile Phe Thr Ser Pro Met				50
835	840	845		
Asp Ile Lys Val Ala Phe Leu Lys Gly Leu Asn Gly Asn Ala Glu Glu				55
850	855	860		
Leu Thr Phe Ser Thr Lys Ser Glu Leu Leu Val Asn Gln Leu Ile Leu				60
865	870	875	880	
Leu Leu Asn Ser Ile Gly Val Ser Asp Ile Lys Ile Glu His Glu Lys				
885	890	895		

DE 198 40 771 A 1

Gly Val Tyr Arg Val Tyr Ile Asn Lys Lys Glu Ser Ser Asn Gly Asp  
900 905 910

<sup>5</sup> Ile Val Leu Asp Ser Val Glu Ser Ile Glu Val Glu Lys Tyr Glu Gly  
915 920 925

<sup>10</sup> Tyr Val Tyr Asp Leu Ser Val Glu Asp Asn Glu Asn Phe Leu Val Gly  
930 935 940

Phe Gly Leu Leu Tyr Ala His Asn Ser Tyr Tyr Gly Tyr Gly Tyr  
<sup>15</sup> 945 950 955 960

Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser Val Thr Ala  
965 970 975

<sup>20</sup> Trp Gly Arg Gln Tyr Ile Asp Leu Val Arg Arg Glu Leu Glu Ala Arg  
980 985 990

<sup>25</sup> Gly Phe Lys Val Leu Tyr Ile Asp Thr Asp Gly Leu Tyr Ala Thr Ile  
995 1000 1005

<sup>30</sup> Pro Gly Val Lys Asp Trp Glu Glu Val Lys Arg Arg Ala Leu Glu Phe  
1010 1015 1020

Val Asp Tyr Ile Asn Ser Lys Leu Pro Gly Val Leu Glu Leu Glu Tyr  
<sup>35</sup> 1025 1030 1035 1040

Glu Gly Phe Tyr Ala Arg Gly Phe Phe Val Thr Lys Lys Tyr Ala  
1045 1050 1055

<sup>40</sup> Leu Ile Asp Glu Glu Gly Lys Ile Val Thr Arg Gly Leu Glu Ile Val  
1060 1065 1070

<sup>45</sup> Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala Arg Val Leu  
1075 1080 1085

Glu Ala Ile Leu Lys His Gly Asn Val Glu Glu Ala Val Lys Ile Val  
<sup>50</sup> 1090 1095 1100

Lys Asp Val Thr Glu Lys Leu Thr Asn Tyr Glu Val Pro Pro Glu Lys  
<sup>55</sup> 1105 1110 1115 1120

Leu Val Ile Tyr Glu Gln Ile Thr Arg Pro Ile Asn Glu Tyr Lys Ala  
1125 1130 1135

<sup>60</sup> Ile Gly Pro His Val Ala Val Ala Lys Arg Leu Met Ala Arg Gly Ile  
1140 1145 1150

DE 198 40 771 A 1

Lys Val Lys Pro Gly Met Val Ile Gly Tyr Ile Val Leu Arg Gly Asp  
1155 1160 1165

Gly Pro Ile Ser Lys Arg Ala Ile Ser Ile Glu Glu Phe Asp Pro Arg  
1170 1175 1180

5

Lys His Lys Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln Val Leu Pro  
1185 1190 1195 1200

10

Ala Val Glu Arg Ile Leu Lys Ala Phe Gly Tyr Lys Arg Glu Asp Leu  
1205 1210 1215

15

Arg Trp Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp Ile Lys Val  
1220 1225 1230

20

Lys Lys Ser  
1235

25

(2) ANGABEN ZU SEQ ID NO: 26:

(i) SEQUENZKENNZEICHEN:  
(A) LÄNGE: 586 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

30

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: *Methanobacterium thermoautotrophicum*

35

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

45

Met Glu Asp Tyr Arg Met Val Leu Leu Asp Ile Asp Tyr Val Thr Val  
1 5 10 15

50

Asp Glu Val Pro Val Ile Arg Leu Phe Gly Lys Asp Lys Ser Gly Gly  
20 25 30

55

Asn Glu Pro Ile Ile Ala His Asp Arg Ser Phe Arg Pro Tyr Ile Tyr  
35 40 45

60

Ala Ile Pro Thr Asp Leu Asp Glu Cys Leu Arg Glu Leu Glu Leu  
50 55 60

65

DE 198 40 771 A 1

Glu Leu Glu Lys Leu Glu Val Lys Glu Met Arg Asp Leu Gly Arg Pro  
65 70 75 80

5 Thr Glu Val Ile Arg Ile Glu Phe Arg His Pro Gln Asp Val Pro Lys  
85 90 95

10 Ile Arg Asp Arg Ile Arg Asp Leu Glu Ser Val Arg Asp Ile Arg Glu  
100 105 110

15 His Asp Ile Pro Phe Tyr Arg Arg Tyr Leu Ile Asp Lys Ser Ile Val  
115 120 125

20 Pro Met Glu Glu Leu Glu Phe Gln Gly Val Glu Val Asp Ser Ala Pro  
130 135 140

25 Ser Val Thr Thr Asp Val Arg Thr Val Glu Val Thr Gly Arg Val Gln  
145 150 155 160

30 Ser Thr Gly Ser Gly Ala His Gly Leu Asp Ile Leu Ser Phe Asp Ile  
165 170 175

35 Glu Val Arg Asn Pro His Gly Met Pro Asp Pro Glu Lys Asp Glu Ile  
180 185 190

40 Val Met Ile Gly Val Ala Gly Asn Met Gly Tyr Glu Ser Val Ile Ser  
195 200 205

45 Thr Ala Gly Asp His Leu Asp Phe Val Glu Val Val Glu Asp Glu Arg  
210 215 220

50 Glu Leu Leu Glu Arg Phe Ala Glu Ile Val Ile Asp Lys Lys Pro Asp  
225 230 235 240

55 Ile Leu Val Gly Tyr Asn Ser Asp Asn Phe Asp Pro Tyr Ile Thr  
245 250 255

60 Arg Arg Ala Ala Ile Leu Gly Ala Glu Leu Asp Leu Gly Trp Asp Gly  
260 265 270

65 Ser Lys Ile Arg Thr Met Arg Arg Gly Phe Ala Asn Ala Thr Ala Ile  
275 280 285

70 Lys Gly Thr Val His Val Asp Leu Tyr Pro Val Met Arg Arg Tyr Met  
290 295 300

75 Asn Leu Asp Arg Tyr Thr Leu Glu Arg Val Tyr Gln Glu Leu Phe Gly  
305 310 315 320

DE 198 40 771 A 1

Glu	Glu	Lys	Ile	Asp	Leu	Pro	Gly	Asp	Arg	Leu	Trp	Glu	Tyr	Trp	Asp	
325																5
Arg	Asp	Glu	Leu	Arg	Asp	Glu	Leu	Phe	Arg	Tyr	Ser	Leu	Asp	Asp	Val	
340																10
Val	Ala	Thr	His	Ile	Ala	Glu	Lys	Ile	Leu	Pro	Leu	Asn	Leu	Glu		
355																15
Leu	Thr	Arg	Leu	Val	Gly	Gln	Pro	Leu	Phe	Asp	Ile	Ser	Arg	Met	Ala	
370																20
Thr	Gly	Gln	Gln	Ala	Glu	Trp	Phe	Leu	Val	Arg	Lys	Ala	Tyr	Gln	Tyr	
385																25
Gly	Glu	Leu	Val	Pro	Asn	Lys	Pro	Ser	Gln	Ser	Asp	Phe	Ser	Ser	Arg	
405																30
Arg	Gly	Arg	Arg	Ala	Val	Gly	Gly	Tyr	Val	Lys	Glu	Pro	Glu	Lys	Gly	
420																35
Leu	His	Glu	Asn	Ile	Val	Gln	Phe	Asp	Phe	Arg	Ser	Leu	Tyr	Pro	Ser	
435																40
Ile	Ile	Ile	Ser	Lys	Asn	Ile	Ser	Pro	Asp	Thr	Leu	Thr	Asp	Asp	Glu	
450																45
Glu	Ser	Glu	Cys	Tyr	Val	Ala	Pro	Glu	Tyr	Gly	Tyr	Arg	Phe	Arg	Lys	
465																50
Ser	Pro	Arg	Gly	Phe	Val	Pro	Ser	Val	Ile	Gly	Glu	Ile	Leu	Ser	Glu	
485																55
Arg	Val	Arg	Ile	Lys	Glu	Glu	Met	Lys	Gly	Ser	Asp	Asp	Pro	Met	Glu	
500																60
Arg	Lys	Ile	Leu	Asn	Val	Gln	Gln	Glu	Ala	Leu	Lys	Arg	Leu	Ala	Asn	
515																65
Thr	Met	Tyr	Gly	Val	Tyr	Gly	Tyr	Ser	Arg	Phe	Arg	Trp	Tyr	Ser	Met	
530																70
Glu	Cys	Ala	Glu	Ala	Ile	Thr	Ala	Trp	Gly	Arg	Asp	Tyr	Ile	Lys	Lys	
545																75
Thr	Ile	Lys	Thr	Ala	Glu	Glu	Phe	Gly	Phe	His	Thr	Val	Tyr	Ala	Asp	
565																80

Thr Asp Gly Phe Tyr Ala Thr Tyr Arg Gly  
580 585

<sup>5</sup> (2) ANGABEN ZU SEQ ID NO: 27:

**(i) SEQUENZKENNZEICHEN**

10 (A) LÄNGE: 1143 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

### 15 (iii) ART DES MOLEKÜLS: Protein

20 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: *Archaeoglobus fulgidus*

**25 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:**

Met Asp Ala Thr Leu Asp Arg Phe Phe Pro Leu Phe Glu Ser Glu Ser  
1 5 10 15

Asn Glu Asp Phe Trp Arg Ile Glu Glu Ile Arg Arg Tyr His Glu Ser  
20 25 30

35 Leu Met Val Glu Leu Asp Arg Ile Tyr Arg Ile Ala Glu Ala Ala Arg  
35 40 45

40 Lys Lys Gly Leu Asp Pro Glu Leu Ser Val Glu Ile Pro Ile Ala Lys  
50 55 60

Asn Met Ala Glu Arg Val Glu Lys Leu Met Asn Leu Gin Gly Leu Ala  
65 70 75 80

Lys Arg Ile Met Glu Leu Glu Glu Gly Gly Leu Ser Arg Glu Leu Ile  
85 90 95

Cys Phe Lys Val Ala Asp Glu Ile Val Glu Gly Lys Phe Gly Glu Met  
100 105 110

Pro Lys Glu Glu Ala Ile Asp Lys Ala Val Arg Thr Ala Val Ala Ile  
115 120 125

60 Met Thr Glu Gly Val Val Ala Ala Pro Ile Glu Gly Ile Ala Arg Val  
130 135 140

## DE 198 40 771 A 1

Arg Ile Asp Arg Glu Asn Phe Leu Arg Val Tyr Tyr Ala Gly Pro Ile			
145	150	155	160
Arg Ser Ala Gly Gly Thr Ala Gln Val Ile Ser Val Val Ala Asp			5
165	170	175	
Tyr Val Arg Arg Lys Ala Glu Ile Gly Arg Tyr Val Pro Thr Glu Glu			10
180	185	190	
Glu Ile Leu Arg Tyr Cys Glu Glu Ile Pro Leu Tyr Lys Lys Val Ala			15
195	200	205	
Asn Leu Gln Tyr Leu Pro Ser Asp Glu Glu Ile Arg Leu Ile Val Ser			
210	215	220	20
Asn Cys Pro Ile Cys Ile Asp Gly Glu Pro Thr Glu Ser Ala Glu Val			
225	230	235	240
Ser Gly Tyr Arg Asn Leu Pro Arg Val Glu Thr Asn Arg Val Arg Gly			25
245	250	255	
Gly Met Ala Leu Val Ile Ala Glu Gly Ile Ala Leu Lys Ala Pro Lys			30
260	265	270	
Leu Lys Lys Met Val Asp Glu Val Gly Ile Glu Gly Trp Glu Trp Leu			35
275	280	285	
Asp Ala Leu Ile Lys Gly Gly Asp Ser Gly Ser Glu Glu Lys			
290	295	300	40
Ala Val Ile Lys Pro Lys Asp Lys Tyr Leu Ser Asp Ile Val Ala Gly			
305	310	315	320
Arg Pro Val Leu Ser His Pro Ser Arg Lys Gly Gly Phe Arg Leu Arg			45
325	330	335	
Tyr Gly Arg Ala Arg Asn Ser Gly Phe Ala Thr Val Gly Val Asn Pro			50
340	345	350	
Ala Thr Met Tyr Leu Leu Glu Phe Val Ala Val Gly Thr Gln Leu Lys			
355	360	365	55
Val Glu Arg Pro Gly Lys Ala Gly Gly Val Val Pro Val Ser Thr Ile			
370	375	380	
Glu Gly Pro Thr Val Arg Leu Lys Asn Gly Asp Val Val Lys Ile Asn			60
385	390	395	400

65

DE 198 40 771 A 1

Thr Leu Ser Glu Ala Lys Ala Leu Lys Gly Glu Val Ala Ala Ile Leu  
405 410 415

5 Asp Leu Gly Glu Ile Leu Ile Asn Tyr Gly Asp Phe Leu Glu Asn Asn  
420 425 430

10 His Pro Leu Ile Pro Ala Ser Tyr Thr Tyr Glu Trp Trp Ile Gln Glu  
435 440 445

15 Ala Glu Lys Ala Gly Leu Arg Gly Asp Tyr Arg Lys Ile Ser Glu Glu  
450 455 460

20 Glu Ala Leu Lys Leu Cys Asp Glu Phe His Val Pro Leu His Pro Asp  
465 470 475 480

25 Tyr Thr Tyr Leu Trp His Asp Ile Ser Val Glu Asp Tyr Arg Tyr Leu  
485 490 495

30 Arg Asn Phe Val Ser Asp Asn Gly Lys Ile Glu Gly Lys His Gly Lys  
500 505 510

35 Ser Val Leu Leu Leu Pro Tyr Asp Ser Arg Val Lys Glu Ile Leu Glu  
515 520 525

40 Ala Leu Leu Leu Glu His Lys Val Arg Glu Ser Phe Ile Val Ile Glu  
530 535 540

45 Thr Trp Arg Ala Phe Ile Arg Cys Leu Gly Leu Asp Glu Lys Leu Ser  
545 550 555 560

50 Lys Val Ser Glu Val Ser Gly Lys Asp Val Leu Glu Ile Val Asn Gly  
565 570 575

55 Ile Ser Gly Ile Lys Val Arg Pro Lys Ala Leu Ser Arg Ile Gly Ala  
580 585 590

60 Arg Met Gly Arg Pro Glu Lys Ala Lys Glu Arg Lys Met Ser Pro Pro  
595 600 605

65 Pro His Ile Leu Phe Pro Val Gly Met Ala Gly Gly Asn Thr Arg Asp  
610 615 620

70 Ile Lys Asn Ala Ile Asn Tyr Thr Lys Ser Tyr Asn Ala Lys Lys Gly  
625 630 635 640

75 Glu Ile Glu Val Glu Ile Ala Ile Arg Lys Cys Pro Gln Cys Gly Lys  
645 650 655

DE 198 40 771 A 1

Glu Thr Phe Trp Leu Lys Cys Asp Val Cys Gly Glu Leu Thr Glu Gln				
660	665	670		
Leu Tyr Tyr Cys Pro Ser Cys Arg Met Lys Asn Thr Ser Ser Val Cys				5
675	680	685		
Glu Ser Cys Gly Arg Glu Cys Glu Gly Tyr Met Lys Arg Lys Val Asp				10
690	695	700		
Leu Arg Glu Leu Tyr Glu Glu Ala Ile Ala Asn Leu Gly Glu Tyr Asp				
705	710	715	720	15
Ser Phe Asp Thr Ile Lys Gly Val Lys Gly Met Thr Ser Lys Thr Lys				
725	730	735		20
Ile Pro Glu Arg Leu Glu Lys Gly Ile Leu Arg Val Lys His Gly Val				
740	745	750		
Phe Val Phe Lys Asp Gly Thr Ala Arg Phe Asp Ala Thr Asp Leu Pro				25
755	760	765		
Ile Thr His Phe Lys Pro Ala Glu Ile Gly Val Ser Val Glu Lys Leu				
770	775	780		30
Arg Glu Leu Gly Tyr Glu Arg Asp Tyr Lys Gly Ala Glu Leu Lys Asn				
785	790	795	800	35
Glu Asn Gln Ile Val Glu Leu Lys Pro Gln Asp Val Ile Leu Pro Lys				
805	810	815		
Ser Gly Ala Glu Tyr Leu Leu Arg Val Ala Asn Phe Ile Asp Asp Leu				40
820	825	830		
Leu Val Lys Phe Tyr Lys Met Glu Pro Phe Tyr Asn Ala Lys Ser Val				
835	840	845		45
Glu Asp Leu Ile Gly His Leu Val Ile Gly Leu Ala Pro His Thr Ser				
850	855	860		50
Ala Gly Val Leu Gly Arg Ile Ile Gly Phe Ser Asp Val Leu Ala Gly				
865	870	875	880	55
Tyr Ala His Pro Tyr Phe His Ala Ala Lys Arg Arg Asn Cys Asp Gly				
885	890	895		
Asp Glu Asp Cys Phe Met Leu Leu Asp Gly Leu Leu Asn Phe Ser				60
900	905	910		

DE 198 40 771 A 1

Arg Lys Phe Leu Pro Asp Lys Arg Gly Gly Gln Met Asp Ala Pro Leu  
915 920 925

5 Val Leu Thr Ala Ile Val Asp Pro Arg Glu Val Asp Lys Glu Val His  
930 935 940

10 Asn Met Asp Ile Val Glu Arg Tyr Pro Leu Glu Phe Tyr Glu Ala Thr  
945 950 955 960

15 Met Arg Phe Ala Ser Pro Lys Glu Met Glu Asp Tyr Val Glu Lys Val  
965 970 975

20 Lys Asp Arg Leu Lys Asp Glu Ser Arg Phe Cys Gly Leu Phe Phe Thr  
980 985 990

25 His Asp Thr Glu Asn Ile Ala Ala Gly Val Lys Glu Ser Ala Tyr Lys  
995 1000 1005

30 Ser Leu Lys Thr Met Gln Asp Lys Val Tyr Arg Gln Met Glu Leu Ala  
1010 1015 1020

35 Arg Met Ile Val Ala Val Asp Glu His Asp Val Ala Glu Arg Val Ile  
1025 1030 1035 1040

40 Asn Val His Phe Leu Pro Asp Ile Ile Gly Asn Leu Arg Ala Phe Ser  
1045 1050 1055

45 Arg Gln Glu Phe Arg Cys Thr Arg Cys Asn Thr Lys Tyr Arg Arg Ile  
1060 1065 1070

50 Pro Leu Val Gly Lys Cys Leu Lys Cys Gly Asn Lys Leu Thr Leu Thr  
1075 1080 1085

55 Val His Ser Ser Ser Ile Met Lys Tyr Leu Glu Leu Ser Lys Phe Leu  
1090 1095 1100

60 Cys Glu Asn Phe Asn Val Ser Ser Tyr Thr Lys Gln Arg Leu Met Leu  
1105 1110 1115 1120

65 Leu Glu Gln Glu Ile Lys Ser Met Phe Glu Asn Gly Thr Glu Lys Gln  
1125 1130 1135

Val Ser Ile Ser Asp Phe Val  
1140

60 (2) ANGABEN ZU SEQ ID NO: 28:

65 (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1139 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

5

(ii) ART DES MOLEKÜLS: Protein

10

(vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Methanococcus jannaschii

15

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

Met Ile Val Met Val His Val Ala Cys Ser Glu Asn Met Lys Lys Tyr  
 1            5            10            15

20

Phe Glu Asn Ile Val Asp Glu Val Lys Lys Ile Tyr Arg Ile Ala Glu  
 20            25            30

25

Glu Cys Arg Lys Lys Gly Phe Asp Pro Thr Asp Glu Val Glu Ile Pro  
 35            40            45

30

Leu Ala Ala Asp Met Ala Asp Arg Val Glu Gly Leu Val Gly Pro Lys  
 50            55            60

35

Gly Val Ala Glu Arg Ile Arg Glu Leu Val Lys Glu Leu Gly Lys Glu  
 65            70            75            80

Pro Ala Ala Leu Glu Ile Ala Lys Glu Ile Val Glu Gly Lys Phe Gly  
 85            90            95

40

Asn Phe Asp Lys Glu Lys Lys Ala Glu Gln Ala Val Arg Thr Ala Leu  
 100            105            110

45

Ala Val Leu Thr Glu Gly Ile Val Ala Ala Pro Leu Glu Gly Ile Ala  
 115            120            125

50

Asp Val Lys Ile Lys Lys Asn Pro Asp Gly Thr Glu Tyr Leu Ala Ile  
 130            135            140

55

Tyr Tyr Ala Gly Pro Ile Arg Ser Ala Gly Gly Thr Ala Gln Ala Leu  
 145            150            155            160

Ser Val Leu Val Gly Asp Phe Val Arg Lys Ala Met Gly Leu Asp Arg  
 165            170            175

60

65

DE 198 40 771 A 1

Tyr Lys Pro Thr Glu Asp Glu Ile Glu Arg Tyr Val Glu Val Glu  
180 185 190

5 Leu Tyr Gln Ser Glu Val Gly Ser Phe Gln Tyr Asn Pro Thr Ala Asp  
195 200 205

10 Glu Ile Arg Thr Ala Ile Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu  
210 215 220

15 Ala Thr Asp Asp Val Glu Val Ser Gly His Arg Asp Leu Pro Arg Val  
225 230 235 240

Glu Thr Asn Gln Leu Arg Gly Gly Ala Leu Leu Val Leu Val Glu Gly  
245 250 255

20 Val Leu Leu Lys Ala Pro Lys Ile Leu Arg His Val Asp Lys Leu Gly  
260 265 270

25 Ile Glu Gly Trp Asp Trp Leu Lys Asp Leu Met Ser Lys Lys Glu Glu  
275 280 285

30 Lys Glu Glu Glu Lys Asp Glu Lys Val Asp Asp Glu Glu Ile Asp Glu  
290 295 300

35 Glu Glu Glu Glu Ile Ser Gly Tyr Trp Arg Asp Val Lys Ile Glu Ala  
305 310 315 320

Asn Lys Lys Phe Ile Ser Glu Val Ile Ala Gly Arg Pro Val Phe Ala  
325 330 335

40 His Pro Ser Lys Val Gly Gly Phe Arg Leu Arg Tyr Gly Arg Ser Arg  
340 345 350

45 Asn Thr Gly Phe Ala Thr Gln Gly Phe His Pro Ala Leu Met Tyr Leu  
355 360 365

50 Val Asp Glu Phe Met Ala Val Gly Thr Gln Leu Lys Thr Glu Arg Pro  
370 375 380

55 Gly Lys Ala Thr Cys Val Val Pro Val Asp Ser Ile Glu Pro Pro Ile  
385 390 395 400

Val Lys Leu Lys Asn Gly Asp Val Ile Arg Val Asp Thr Ile Glu Lys  
405 410 415

60 Ala Met Asp Val Arg Asn Arg Val Glu Glu Ile Leu Phe Leu Gly Asp  
420 425 430

DE 198 40 771 A 1

Val	Leu	Val	Asn	Tyr	Gly	Asp	Phe	Leu	Glu	Asn	Asn	His	Pro	Leu	Leu		
435																	
Pro	Ser	Cys	Trp	Cys	Glu	Glu	Trp	Tyr	Glu	Ile	Leu	Ile	Ala	Asn		5	
450																	
Asn	Ile	Glu	Tyr	Asp	Lys	Asp	Phe	Ile	Lys	Asn	Pro	Lys	Pro	Glu		10	
465																	
Ala	Val	Lys	Phe	Ala	Leu	Glu	Thr	Lys	Thr	Pro	Leu	His	Pro	Arg	Phe		15
485																	
Thr	Tyr	His	Trp	His	Asp	Val	Ser	Lys	Glu	Asp	Ile	Ile	Leu	Leu	Arg		20
500																	
Asn	Trp	Leu	Leu	Lys	Gly	Lys	Glu	Asp	Ser	Leu	Glu	Gly	Lys	Lys	Val		25
515																	
Trp	Ile	Val	Asp	Leu	Glu	Ile	Glu	Glu	Asp	Lys	Lys	Ala	Lys	Arg	Ile		30
530																	
Leu	Glu	Leu	Ile	Gly	Cys	Cys	His	Leu	Val	Arg	Asn	Lys	Lys	Val	Ile		35
545																	
Ile	Glu	Glu	Tyr	Tyr	Pro	Leu	Leu	Tyr	Ser	Leu	Gly	Phe	Asp	Val	Glu		40
565																	
Asn	Lys	Lys	Asp	Leu	Val	Glu	Asn	Ile	Glu	Lys	Ile	Leu	Glu	Ser	Ala		45
580																	
Lys	Asn	Ser	Met	His	Leu	Ile	Asn	Leu	Leu	Ala	Pro	Phe	Glu	Val	Arg		50
595																	
Arg	Asn	Thr	Tyr	Val	Tyr	Val	Gly	Ala	Arg	Met	Gly	Arg	Pro	Glu	Lys		55
610																	
Ala	Ala	Pro	Arg	Lys	Met	Lys	Pro	Pro	Val	Asn	Gly	Leu	Phe	Pro	Ile		60
625																	
Gly	Asn	Ala	Gly	Gly	Gln	Val	Arg	Leu	Ile	Asn	Lys	Ala	Val	Glu	Glu		65
645																	
Asn	Asn	Thr	Asp	Asp	Val	Asp	Val	Ser	Tyr	Thr	Arg	Cys	Pro	Asn	Cys		70
660																	
Gly	Lys	Ile	Ser	Leu	Tyr	Arg	Val	Cys	Pro	Phe	Cys	Gly	Thr	Lys	Val		75
675																	

DE 198 40 771 A 1

Glu Leu Asp Asn Phe Gly Arg Ile Lys Ala Pro Leu Lys Asp Tyr Trp  
 690 695 700

<sup>5</sup> Tyr Ala Ala Leu Lys Arg Leu Gly Ile Asn Lys Pro Gly Asp Val Lys  
 705 710 715 720

<sup>10</sup> Cys Ile Lys Gly Met Thr Ser Lys Gln Lys Ile Val Glu Pro Leu Glu  
 725 730 735

<sup>15</sup> Lys Ala Ile Leu Arg Ala Ile Asn Glu Val Tyr Val Phe Lys Asp Gly  
 740 745 750

<sup>20</sup> Thr Thr Arg Phe Asp Cys Thr Asp Val Pro Val Thr His Phe Lys Pro  
 755 760 765

Asn Glu Ile Asn Val Thr Val Glu Lys Leu Arg Glu Leu Gly Tyr Asp  
 770 775 780

<sup>25</sup> Lys Asp Ile Tyr Gly Asn Glu Leu Val Asp Gly Glu Gln Val Val Glu  
 785 790 795 800

<sup>30</sup> Leu Lys Pro Gln Asp Val Ile Ile Pro Glu Ser Cys Ala Glu Tyr Phe  
 805 810 815

<sup>35</sup> Val Lys Val Ala Asn Phe Ile Asp Asp Leu Leu Glu Lys Phe Tyr Lys  
 820 825 830

Val Glu Arg Phe Tyr Asn Val Lys Lys Glu Asp Leu Ile Gly His  
 835 840 845

<sup>40</sup> Leu Val Ile Gly Met Ala Pro His Thr Ser Ala Gly Met Val Gly Arg  
 850 855 860

<sup>45</sup> Ile Ile Gly Tyr Thr Lys Ala Asn Val Gly Tyr Ala His Pro Tyr Phe  
 865 870 875 880

<sup>50</sup> His Ala Ala Lys Arg Arg Asn Cys Asp Gly Asp Glu Asp Ser Phe Phe  
 885 890 895

<sup>55</sup> Leu Leu Leu Asp Ala Phe Leu Asn Phe Ser Lys Lys Phe Leu Pro Asp  
 900 905 910

Lys Arg Gly Gly Gln Met Asp Ala Pro Leu Val Leu Thr Thr Ile Leu  
 915 920 925

<sup>60</sup> Asp Pro Lys Glu Val Asp Gly Glu Val His Asn Met Asp Thr Met Trp  
 930 935 940

DE 198 40 771 A 1

Ser Tyr Pro Leu Glu Phe Tyr Glu Lys Thr Leu Glu Met Pro Ser Pro				
945	950	955	960	
Lys Glu Val Lys Glu Phe Met Glu Thr Val Glu Asp Arg Leu Gly Lys				5
	965	970	975	
Pro Glu Gln Tyr Glu Gly Ile Gly Tyr Thr His Glu Thr Ser Arg Ile				10
980	985	990		
Asp Leu Gly Pro Lys Val Cys Ala Tyr Lys Thr Leu Gly Ser Met Leu				15
995	1000	1005		
Glu Lys Thr Thr Ser Gln Leu Ser Val Ala Lys Lys Ile Arg Ala Thr				20
1010	1015	1020		
Asp Glu Arg Asp Val Ala Glu Lys Val Ile Gln Ser His Phe Ile Pro				25
1025	1030	1035	1040	
Asp Leu Ile Gly Asn Leu Arg Ala Phe Ser Arg Gln Ala Val Arg Cys				30
1045	1050	1055		
Lys Cys Gly Ala Lys Tyr Arg Arg Ile Pro Leu Lys Gly Lys Cys Pro				35
1060	1065	1070		
Lys Cys Gly Ser Asn Leu Ile Leu Thr Val Ser Lys Gly Ala Val Glu				40
1075	1080	1085		
Lys Tyr Met Asp Val Ala Glu Lys Met Ala Glu Glu Tyr Asn Val Asn				45
1090	1095	1100		
Asp Tyr Ile Lys Gln Arg Leu Lys Ile Ile Lys Glu Gly Ile Asn Ser				50
1105	1110	1115	1120	
Ile Phe Glu Asn Glu Lys Ser Arg Gln Val Lys Leu Ser Asp Phe Phe				55
1125	1130	1135		
Lys Ile Gly				60

(2) ANGABEN ZU SEQ ID NO: 29:

- (i) SEQUENZKENNZEICHEN:
  - (A) LÄNGE: 1434 Aminosäuren
  - (B) ART: Aminosäure
  - (C) STRANGFORM: Einzelstrang
  - (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: Protein

65

## (vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: *Pyrococcus horikoshii*

5

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

10	Met Val Leu Met Glu Leu Pro Lys Glu Met Glu Glu Tyr Phe Ser Met			
	1	5	10	15
15	Leu Gln Arg Glu Ile Asp Lys Ala Tyr Glu Ile Ala Lys Lys Ala Arg			
	20	25	30	
20	Ala Gln Gly Lys Asp Pro Ser Leu Asp Val Glu Ile Pro Gln Ala Ser			
	35	40	45	
25	Asp Met Ala Gly Arg Val Glu Ser Leu Val Gly Pro Pro Gly Val Ala			
	50	55	60	
30	Glu Arg Ile Arg Glu Leu Val Lys Glu Tyr Gly Lys Glu Ile Ala Ala			
	65	70	75	80
35	Leu Lys Ile Val Asp Glu Ile Ile Asp Gly Lys Phe Gly Asp Leu Gly			
	85	90	95	
40	Ser Lys Glu Lys Tyr Ala Glu Gln Ala Val Arg Thr Ala Leu Ala Ile			
	100	105	110	
45	Leu Thr Glu Gly Val Val Ser Ala Pro Ile Glu Gly Ile Ala Ser Val			
	115	120	125	
50	Lys Ile Lys Arg Asn Thr Trp Ser Asp Asn Ser Glu Tyr Leu Ala Leu			
	130	135	140	
55	Tyr Tyr Ala Gly Pro Ile Arg Ser Ser Gly Gly Thr Ala Gln Ala Leu			
	145	150	155	160
60	Ser Val Leu Val Gly Asp Tyr Val Arg Arg Lys Leu Gly Leu Asp Arg			
	165	170	175	
65	Phe Lys Pro Ser Glu Lys His Ile Glu Arg Met Val Glu Glu Val Asp			
	180	185	190	
70	Leu Tyr His Arg Thr Val Ser Arg Leu Gln Tyr His Pro Ser Pro Glu			
	195	200	205	
75	Glu Val Arg Leu Ala Met Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu			
	210	215	220	

65

DE 198 40 771 A 1

Ala Thr Asp Glu Val Glu Val Ser His Arg Asp Ile Pro Gly Val Glu				
225	230	235	240	
Thr Asn Gln Leu Arg Gly Gly Ala Ile Leu Val Leu Ala Glu Gly Val				5
245	250	255		
Leu Gln Lys Ala Lys Lys Leu Val Lys Tyr Ile Asp Lys Met Gly Ile				10
260	265	270		
Glu Gly Trp Glu Trp Leu Lys Glu Phe Val Glu Ala Lys Glu Lys Gly				15
275	280	285		
Glu Glu Ile Glu. Glu Glu Gly Ser Ala Glu Ser Thr Val Glu Glu Thr				
290	295	300		
Lys Val Glu Val Asp Met Gly Phe Tyr Tyr Ser Leu Tyr Gln Lys Phe				20
305	310	315	320	
Lys Ser Glu Ile Ala Pro Asn Asp Lys Tyr Ala Lys Glu Ile Ile Gly				25
325	330	335		
Gly Arg Pro Leu Phe Ser Asp Pro Ser Arg Asn Gly Gly Phe Arg Leu				30
340	345	350		
Arg Tyr Gly Arg Ser Arg Val Ser Gly Phe Ala Thr Trp Gly Ile Asn				35
355	360	365		
Pro Ala Thr Met Ile Leu Val Asp Glu Phe Leu Ala Ile Gly Thr Gln				
370	375	380		
Leu Lys Thr Glu Arg Pro Gly Lys Gly Ala Val Val Thr Pro Val Thr				40
385	390	395	400	
Thr Ile Glu Gly Pro Ile Val Lys Leu Lys Asp Gly Ser Val Val Lys				45
405	410	415		
Val Asp Asp Tyr Lys Leu Ala Leu Lys Ile Arg Asp Glu Val Glu Glu				50
420	425	430		
Ile Leu Tyr Leu Gly Asp Ala Val Ile Ala Phe Gly Asp Phe Val Glu				
435	440	445		
Asn Asn Gln Thr Leu Leu Pro Ala Asn Tyr Cys Glu Glu Trp Trp Ile				55
450	455	460		
Leu Glu Phe Thr Lys Ala Leu Asn Glu Ile Tyr Glu Val Glu Leu Lys				60
465	470	475	480	

65

## DE 198 40 771 A 1

Pro Phe Glu Val Asn Ser Ser Glu Asp Leu Glu Glu Ala Ala Asp Tyr  
 485 490 495

5 Leu Glu Val Asp Ile Glu Phe Leu Lys Glu Leu Leu Lys Asp Pro Leu  
 500 505 510

10 Arg Thr Lys Pro Pro Val Glu Leu Ala Ile His Phe Ser Glu Ile Leu  
 515 520 525

15 Gly Ile Pro Leu His Pro Tyr Tyr Thr Leu Tyr Trp Asn Ser Val Lys  
 530 535 540

20 Pro Glu Gln Val Glu Lys Leu Trp Arg Val Leu Lys Glu His Ala His  
 545 550 555 560

Ile Asp Trp Asp Asn Phe Arg Gly Ile Lys Phe Ala Arg Arg Ile Val  
 565 570 575

25 Ile Pro Leu Glu Lys Leu Arg Asp Ser Lys Arg Ala Leu Glu Leu Leu  
 580 585 590

30 Gly Leu Pro His Lys Val Glu Gly Lys Asn Val Ile Val Asp Tyr Pro  
 595 600 605

35 Trp Ala Ala Ala Leu Leu Thr Pro Leu Gly Asn Leu Glu Trp Glu Phe  
 610 615 620

Arg Ala Lys Pro Leu His Thr Thr Ile Asp Ile Ile Asn Glu Asn Asn  
 625 630 635 640

40 Glu Ile Lys Leu Arg Asp Arg Gly Ile Ser Trp Ile Gly Ala Arg Met  
 645 650 655

45 Gly Arg Pro Glu Lys Ala Lys Glu Arg Lys Met Lys Pro Pro Val Gln  
 660 665 670

50 Val Leu Phe Pro Ile Gly Leu Ala Gly Gly Ser Ser Arg Asp Ile Lys  
 675 680 685

55 Lys Ala Ala Glu Glu Gly Lys Val Ala Glu Val Glu Ile Ala Leu Phe  
 690 695 700

Lys Cys Pro Lys Cys Gly His Val Gly Pro Glu His Ile Cys Pro Asn  
 705 710 715 720

60 Cys Gly Thr Arg Lys Glu Leu Ile Trp Val Cys Pro Arg Cys Asn Ala  
 725 730 735

DE 198 40 771 A 1

Glu	Tyr	Pro	Glu	Ser	Gln	Ala	Ser	Gly	Tyr	Asn	Tyr	Thr	Cys	Pro	Lys		
740																	5
Cys	Asn	Val	Lys	Leu	Lys	Pro	Tyr	Ala	Lys	Arg	Lys	Ile	Lys	Pro	Ser		
755																	10
Glu	Leu	Leu	Lys	Arg	Ala	Met	Asp	Asn	Val	Lys	Val	Tyr	Gly	Ile	Asp		
770																	15
Lys	Leu	Lys	Gly	Val	Met	Gly	Met	Thr	Ser	Gly	Trp	Lys	Met	Pro	Glu		
785																	20
Pro	Leu	Glu	Lys	Gly	Leu	Leu	Arg	Ala	Lys	Asn	Asp	Val	Tyr	Val	Phe		
805																	25
Lys	Asp	Gly	Thr	Ile	Arg	Phe	Asp	Ala	Thr	Asp	Ala	Pro	Ile	Thr	His		
820																	30
Phe	Arg	Pro	Arg	Glu	Ile	Gly	Val	Ser	Val	Glu	Lys	Leu	Arg	Glu	Leu		
835																	35
Gly	Tyr	Thr	His	Asp	Phe	Glu	Gly	Asn	Pro	Leu	Val	Ser	Glu	Asp	Gln		
850																	40
Ile	Val	Glu	Leu	Lys	Pro	Gln	Asp	Ile	Ile	Leu	Ser	Lys	Glu	Ala	Gly		
865																	45
Lys	Tyr	Leu	Leu	Lys	Val	Ala	Lys	Phe	Val	Asp	Asp	Leu	Leu	Glu	Lys		
885																	50
Phe	Tyr	Gly	Leu	Pro	Arg	Phe	Tyr	Asn	Ala	Glu	Lys	Met	Glu	Asp	Leu		
900																	55
Ile	Gly	His	Leu	Val	Ile	Gly	Leu	Ala	Pro	His	Thr	Ser	Ala	Gly	Ile		
915																	60
Val	Gly	Arg	Ile	Ile	Gly	Phe	Val	Asp	Ala	Leu	Val	Gly	Tyr	Ala	His		
930																	65
Pro	Tyr	Phe	His	Ala	Ala	Lys	Arg	Arg	Asn	Cys	Phe	Pro	Gly	Asp	Thr		
945																	70
Arg	Ile	Leu	Val	Gln	Ile	Asn	Gly	Thr	Pro	Gln	Arg	Val	Thr	Leu	Lys		
965																	75
Glu	Leu	Tyr	Glu	Leu	Phe	Asp	Glu	Glu	His	Tyr	Glu	Ser	Met	Val	Tyr		
980																	80

## DE 198 40 771 A 1

Val Arg Lys Lys Pro Lys Val Asp Ile Lys Val Tyr Ser Phe Asn Pro  
 995 1000 1005

5 Glu Glu Gly Lys Val Val Leu Thr Asp Ile Glu Glu Val Ile Lys Ala  
 1010 1015 1020

10 Pro Ala Thr Asp His Leu Ile Arg Phe Glu Leu Glu Leu Gly Ser Ser  
 1025 1030 1035 1040

15 Phe Glu Thr Thr Val Asp His Pro Val Leu Val Tyr Glu Asn Gly Lys  
 1045 1050 1055

20 Phe Val Glu Lys Arg Ala Phe Glu Val Arg Glu Gly Asn Ile Ile Ile  
 1060 1065 1070

25 Ile Ile Asp Glu Ser Thr Leu Glu Pro Leu Lys Val Ala Val Lys Lys  
 1075 1080 1085

30 Ile Glu Phe Ile Glu Pro Pro Glu Asp Phe Val Phe Ser Leu Asn Ala  
 1090 1095 1100

35 Lys Lys Tyr His Thr Val Ile Ile Asn Glu Asn Ile Val Thr His Gln  
 1105 1110 1115 1120

40 Cys Asp Gly Asp Glu Asp Ala Val Met Leu Leu Leu Asp Ala Leu Leu  
 1125 1130 1135

45 Asn Phe Ser Arg Tyr Tyr Leu Pro Glu Lys Arg Gly Gly Lys Met Asp  
 1140 1145 1150

50 Ala Pro Leu Val Ile Thr Thr Arg Leu Asp Pro Arg Glu Val Asp Ser  
 1155 1160 1165

55 Glu Val His Asn Met Asp Ile Val Arg Tyr Tyr Pro Leu Glu Phe Tyr  
 1170 1175 1180

60 Glu Ala Thr Tyr Glu Leu Lys Ser Pro Lys Glu Leu Val Gly Val Ile  
 1185 1190 1195 1200

65 Glu Arg Val Glu Asp Arg Leu Gly Lys Pro Glu Met Tyr Tyr Gly Leu  
 1205 1210 1215

70 Lys Phe Thr His Asp Thr Asp Asp Ile Ala Leu Gly Pro Lys Met Ser  
 1220 1225 1230

75 Leu Tyr Lys Gln Leu Gly Asp Met Glu Glu Lys Val Lys Arg Gln Leu  
 1235 1240 1245

DE 198 40 771 A 1

Asp Val Ala Arg Arg Ile Arg Ala Val Asp Glu His Lys Val Ala Glu				
1250	1255	1260		
5				
Thr Ile Leu Asn Ser His Leu Ile Pro Asp Leu Arg Gly Asn Leu Arg				
1265	1270	1275	1280	
10				
Ser Phe Thr Arg Gln Glu Phe Arg Cys Val Lys Cys Asn Thr Lys Phe				
1285	1290	1295		
15				
Arg Arg Pro Pro Leu Asp Gly Lys Cys Pro Ile Cys Gly Gly Lys Ile				
1300	1305	1310		
20				
Val Leu Thr Val Ser Lys Gly Ala Ile Glu Lys Tyr Leu Gly Thr Ala				
1315	1320	1325		
25				
Lys Met Leu Val Thr Glu Tyr Lys Val Lys Asn Tyr Thr Arg Gln Arg				
1330	1335	1340		
Ile Cys Leu Thr Glu Arg Asp Ile Asp Ser Leu Phe Glu Thr Val Phe				
1345	1350	1355	1360	
30				
Pro Glu Thr Gln Leu Thr Leu Leu Val Asn Pro Asn Asp Ile Cys Gln				
1365	1370	1375		
35				
Arg Ile Ile Met Glu Arg Thr Gly Gly Ser Lys Lys Ser Gly Leu Leu				
1380	1385	1390		
40				
Glu Asn Phe Ala Asn Gly Tyr Asn Lys Gly Lys Lys Glu Met Pro				
1395	1400	1405		
Lys Lys Gin Arg Lys Lys Glu Gln Glu Lys Ser Lys Lys Arg Lys Val				
1410	1415	1420		
Ile Ser Leu Asp Asp Phe Phe Ser Arg Lys				
1425	1430			
45				
(2) ANGABEN ZU SEQ ID NO: 30:				50
55				
(i) SEQUENZKENNZEICHEN:				
(A) LÄNGE: 1092 Aminosäuren				
(B) ART: Aminosäure				
(C) STRANGFORM: Einzelstrang				
(D) TOPOLOGIE: linear				
60				
(ii) ART DES MOLEKÜLS: Protein				
65				
(vi) URSPRÜNLICHE HERKUNFT:				
(A) ORGANISMUS: Methanobacterium thermoautotrophicum				

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

5 Met Met Asp Tyr Phe Asn Glu Leu Glu Arg Glu Thr Glu Arg Leu Tyr  
 1 5 10 15

10 Glu Ile Ala Arg Lys Ala Arg Ala Arg Gly Leu Asp Val Ser Thr Thr  
 20 25 30

15 Pro Glu Ile Pro Leu Ala Lys Asp Leu Ala Glu Arg Val Glu Gly Leu  
 35 40 45

20 Val Gly Pro Glu Gly Ile Ala Arg Arg Ile Lys Glu Leu Glu Gly Asp  
 50 55 60

25 Arg Gly Arg Glu Glu Val Ala Phe Gln Ile Ala Ala Glu Ile Ala Ser  
 65 70 75 80

30 Gln Ala Val Pro Asp Asp Asp Pro Glu Glu Arg Glu Lys Leu Ala Asp  
 85 90 95

35 Gln Ala Leu Arg Thr Ala Leu Ala Ile Leu Thr Glu Gly Val Val Ala  
 100 105 110

40 Ala Pro Leu Glu Gly Ile Ala Arg Val Arg Ile Lys Glu Asn Phe Asp  
 115 120 125

45 Lys Ser Arg Tyr Leu Ala Val Tyr Phe Ala Gly Pro Ile Arg Ser Ala  
 130 135 140

50 Gly Gly Thr Ala Ala Ala Leu Ser Val Leu Ile Ala Asp Tyr Ile Arg  
 145 150 155 160

55 Leu Ala Val Gly Leu Asp Arg Tyr Lys Pro Val Glu Arg Glu Ile Glu  
 165 170 175

60 Arg Tyr Val Glu Glu Val Glu Leu Tyr Glu Ser Glu Val Thr Asn Leu  
 180 185 190

65 Gln Tyr Ser Pro Lys Pro Asp Glu Val Arg Leu Ala Ala Ser Lys Ile  
 195 200 205

70 Pro Val Glu Val Thr Gly Glu Pro Thr Asp Lys Val Glu Val Ser His  
 210 215 220

75 Arg Asp Leu Glu Arg Val Glu Thr Asn Asn Ile Arg Gly Ala Leu  
 225 230 235 240

## DE 198 40 771 A 1

Leu Ala Met Val Glu Gly Val Ile Gln Lys Ala Pro Lys Val Leu Lys			
245	250	255	
Tyr Ala Lys Gln Leu Lys Leu Glu Gly Trp Asp Trp Leu Glu Lys Phe			5
260	265	270	
Ser Lys Ala Pro Lys Lys Gly Glu Gly Glu Lys Val Val Val Lys			10
275	280	285	
Ala Asp Ser Lys Tyr Val Glu Asp Ile Ile Gly Gly Arg Pro Val Leu			15
290	295	300	
Ala Tyr Pro Ser Glu Lys Gly Ala Phe Arg Leu Arg Tyr Gly Arg Ala			20
305	310	315	320
Arg Asn Thr Gly Leu Ala Ala Met Gly Val His Pro Ala Thr Met Glu			25
325	330	335	
Leu Leu Gln Phe Leu Ala Val Gly Thr Gln Met Lys Ile Glu Arg Pro			30
340	345	350	
Gly Lys Gly Asn Cys Val Val Pro Val Asp Thr Ile Asp Gly Pro Val			35
355	360	365	
Val Lys Leu Arg Asn Gly Asp Val Ile Arg Ile Glu Asp Ala Glu Thr			40
370	375	380	
Ala Ser Arg Val Arg Ser Glu Val Glu Glu Ile Leu Phe Leu Gly Asp			45
385	390	395	400
Met Leu Val Ala Phe Gly Glu Phe Leu Arg Asn Asn His Val Leu Met			50
405	410	415	
Pro Ala Gly Trp Cys Glu Glu Trp Trp Ile Gln Thr Ile Leu Ser Ser			55
420	425	430	
Pro Lys Tyr Pro Gly Asp Asp Pro Leu Asn Leu Ser Tyr Tyr Arg Thr			60
435	440	445	
Arg Trp Asn Glu Leu Glu Val Ser Ala Gly Asp Ala Phe Arg Ile Ser			65
450	455	460	
Glu Glu Tyr Asp Val Pro Leu His Pro Arg Tyr Thr Tyr Phe Tyr His			
465	470	475	480
Asp Val Thr Val Arg Glu Leu Asn Met Leu Arg Glu Trp Leu Asn Thr			
485	490	495	

## DE 198 40 771 A 1

Ser Gln Leu Glu Asp Glu Leu Val Leu Glu Leu Arg Pro Glu Lys Arg  
 500 505 510

<sup>5</sup> Ile Leu Glu Ile Leu Gly Val Pro His Arg Val Lys Asp Ser Arg Val  
 515 520 525

<sup>10</sup> Val Ile Gly His Asp Asp Ala His Ala Leu Ile Lys Thr Leu Arg Lys  
 530 535 540

Pro Leu Glu Asp Ser Ser Asp Thr Val Glu Ala Leu Asn Arg Val Ser  
 15 545 550 555 560

Pro Val Arg Ile Met Lys Lys Ala Pro Thr Tyr Ile Gly Thr Arg Val  
 565 570 575

<sup>20</sup> Gly Arg Pro Glu Lys Thr Lys Glu Arg Lys Met Arg Pro Ala Pro His  
 580 585 590

<sup>25</sup> Val Leu Phe Pro Ile Gly Lys Tyr Gly Gly Ser Arg Arg Asn Ile Pro  
 595 600 605

<sup>30</sup> Asp Ala Ala Lys Lys Gly Ser Ile Thr Val Glu Ile Gly Arg Ala Thr  
 610 615 620

Cys Pro Ser Cys Arg Val Ser Ser Met Gln Ser Ile Cys Pro Ser Cys  
 35 625 630 635 640

Gly Ser Arg Thr Val Ile Gly Glu Pro Gly Lys Arg Asn Ile Asn Leu  
 645 650 655

<sup>40</sup> Ala Ala Leu Leu Lys Arg Ala Ala Glu Asn Val Ser Val Arg Lys Leu  
 660 665 670

<sup>45</sup> Asp Glu Ile Lys Gly Val Glu Gly Met Ile Ser Ala Glu Lys Phe Pro  
 675 680 685

Glu Pro Leu Glu Lys Gly Ile Leu Arg Ala Lys Asn Asp Val Tyr Thr  
 50 690 695 700

<sup>55</sup> Phe Lys Asp Ala Thr Ile Arg His Asp Ser Thr Asp Leu Pro Leu Thr  
 705 710 715 720

His Phe Thr Pro Arg Glu Val Gly Val Ser Val Glu Arg Leu Arg Glu  
 725 730 735

<sup>60</sup> Leu Gly Tyr Thr Arg Asp Cys Tyr Gly Asp Glu Leu Glu Asp Glu Asp  
 740 745 750

## DE 198 40 771 A 1

Gln Ile Leu Glu Leu Arg Val Gln Asp Val Val Ile Ser Glu Asp Cys				
755	760	765		
Ala Asp Tyr Leu Val Arg Val Ala Asn Phe Val Asp Asp Leu Leu Glu				5
770	775	780		
Arg Phe Tyr Asp Leu Glu Arg Phe Tyr Asn Val Lys Thr Arg Glu Asp				10
785	790	795	800	
Leu Val Gly His Leu Ile Ala Gly Leu Ala Pro His Thr Ser Ala Ala				
805	810	815		15
Val Leu Gly Arg Ile Ile Gly Phe Thr Gly Ala Ser Ala Cys Tyr Ala				
820	825	830		20
His Pro Tyr Phe His Ser Ala Lys Arg Arg Asn Cys Asp Ser Asp Glu				
835	840	845		
Asp Ser Val Met Leu Leu Leu Asp Ala Leu Leu Asn Phe Ser Lys Ser				25
850	855	860		
Tyr Leu Pro Ser Ser Arg Gly Gly Ser Met Asp Ala Pro Leu Val Leu				
865	870	875	880	30
Ser Thr Arg Ile Asp Pro Glu Glu Ile Asp Asp Glu Ser His Asn Ile				
885	890	895		35
Asp Thr Met Asp Met Ile Pro Leu Glu Val Tyr Glu Arg Ser Phe Asp				
900	905	910		
His Pro Arg Pro Ser Glu Val Leu Asp Val Ile Asp Asn Val Glu Lys				40
915	920	925		
Arg Leu Gly Lys Pro Glu Gln Tyr Thr Gly Leu Met Phe Ser His Asn				
930	935	940		45
Thr Ser Arg Ile Asp Glu Gly Pro Lys Val Cys Leu Tyr Lys Leu Leu				
945	950	955	960	50
Pro Thr Met Lys Glu Lys Val Glu Ser Gln Ile Thr Leu Ala Glu Lys				
965	970	975		
Ile Arg Ala Val Asp Gln Arg Ser Val Val Glu Gly Val Leu Met Ser				
980	985	990		55
His Phe Leu Pro Asp Met Met Gly Asn Ile Arg Ala Phe Ser Arg Gln				
995	1000	1005		60

65

DE 198 40 771 A 1

Lys Val Arg Cys Thr Lys Cys Asn Arg Lys Tyr Arg Arg Ile Pro Leu  
1010 1015 1020

5 Ser Gly Glu Cys Arg Cys Gly Gly Asn Leu Val Leu Thr Val Ser Lys  
1025 1030 1035 1040

10 Gly Ser Val Ile Lys Tyr Leu Glu Ile Ser Lys Glu Leu Ala Ser Arg  
1045 1050 1055

15 Tyr Pro Ile Asp Pro Tyr Leu Met Gln Arg Ile Glu Ile Leu Glu Tyr  
1060 1065 1070

Gly Val Asn Ser Leu Phe Glu Ser Asp Arg Ser Lys Gln Ser Ser Leu  
1075 1080 1085

20 Asp Val Phe Leu  
1090

25 (2) ANGABEN ZU SEQ ID NO: 31:

(i) SEQUENZKENNZEICHEN:

30 (A) LÄNGE: 1263 Aminosäuren  
(B) ART: Aminosäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

35 (ii) ART DES MOLEKÜLS: Protein

40 (vi) URSPRÜNLICHE HERKUNFT:  
(A) ORGANISMUS: Pyrococcus furiosus

45 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

Met Glu Leu Pro Lys Glu Ile Glu Glu Tyr Phe Glu Met Leu Gln Arg  
1 5 10 15

50 Glu Ile Asp Lys Ala Tyr Glu Ile Ala Lys Lys Ala Arg Ser Gln Gly  
20 25 30

55 Lys Asp Pro Ser Thr Asp Val Glu Ile Pro Gln Ala Thr Asp Met Ala  
35 40 45

60 Gly Arg Val Glu Ser Leu Val Gly Pro Pro Gly Val Ala Gln Arg Ile  
50 55 60

65

## DE 198 40 771 A 1

Arg Glu Leu Leu Lys Glu Tyr Asp Lys Glu Ile Val Ala Leu Lys Ile				
65	70	75	80	
Val Asp Glu Ile Ile Glu Gly Lys Phe Gly Asp Phe Gly Ser Lys Glu				5
85	90	95		
Lys Tyr Ala Glu Gln Ala Val Arg Thr Ala Leu Ala Ile Leu Thr Glu				10
100	105	110		
Gly Ile Val Ser Ala Pro Leu Glu Gly Ile Ala Asp Val Lys Ile Lys				15
115	120	125		
Arg Asn Thr Trp Ala Asp Asn Ser Glu Tyr Leu Ala Leu Tyr Tyr Ala				20
130	135	140		
Gly Pro Ile Arg Ser Ser Gly Gly Thr Ala Gln Ala Leu Ser Val Leu				
145	150	155	160	
Val Gly Asp Tyr Val Arg Arg Lys Leu Gly Leu Asp Arg Phe Lys Pro				25
165	170	175		
Ser Gly Lys His Ile Glu Arg Met Val Glu Glu Val Asp Leu Tyr His				30
180	185	190		
Arg Ala Val Ser Arg Leu Gln Tyr His Pro Ser Pro Asp Glu Val Arg				35
195	200	205		
Leu Ala Met Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu Ala Thr Asp				
210	215	220		40
Asp Val Glu Val Ser His Arg Asp Val Glu Gly Val Glu Thr Asn Gln				
225	230	235	240	
Leu Arg Gly Gly Ala Ile Leu Val Leu Ala Glu Gly Val Leu Gln Lys				45
245	250	255		
Ala Lys Lys Leu Val Lys Tyr Ile Asp Lys Met Gly Ile Asp Gly Trp				
260	265	270		50
Glu Trp Leu Lys Glu Phe Val Glu Ala Lys Glu Lys Gly Glu Glu Ile				
275	280	285		
Glu Glu Ser Glu Ser Lys Ala Glu Glu Ser Lys Val Glu Thr Arg Val				
290	295	300		
Glu Val Glu Lys Gly Phe Tyr Tyr Lys Leu Tyr Glu Lys Phe Arg Ala				60
305	310	315	320	

DE 198 40 771 A 1

Glu Ile Ala Pro Ser Glu Lys Tyr Ala Lys Glu Ile Ile Gly Gly Arg  
 325 330 335

<sup>5</sup> Pro Leu Phe Ala Gly Pro Ser Glu Asn Gly Gly Phe Arg Leu Arg Tyr  
 340 345 350

<sup>10</sup> Gly Arg Ser Arg Val Ser Gly Phe Ala Thr Trp Ser Ile Asn Pro Ala  
 355 360 365

Thr Met Val Leu Val Asp Glu Phe Leu Ala Ile Gly Thr Gln Met Lys  
<sup>15</sup> 370 375 380

Thr Glu Arg Pro Gly Lys Gly Ala Val Val Thr Pro Ala Thr Thr Ala  
 385 390 395 400

<sup>20</sup> Glu Gly Pro Ile Val Lys Leu Lys Asp Gly Ser Val Val Arg Val Asp  
 405 410 415

<sup>25</sup> Asp Tyr Asn Leu Ala Leu Lys Ile Arg Asp Glu Val Glu Glu Ile Leu  
 420 425 430

<sup>30</sup> Tyr Leu Gly Asp Ala Ile Ile Ala Phe Gly Asp Phe Val Glu Asn Asn  
 435 440 445

<sup>35</sup> Gln Thr Leu Leu Pro Ala Asn Tyr Val Glu Glu Trp Trp Ile Gln Glu  
 450 455 460

Phe Val Lys Ala Val Asn Glu Ala Tyr Glu Val Glu Leu Arg Pro Phe  
 465 470 475 480

<sup>40</sup> Glu Glu Asn Pro Arg Glu Ser Val Glu Glu Ala Ala Glu Tyr Leu Glu  
 485 490 495

<sup>45</sup> Val Asp Pro Glu Phe Leu Ala Lys Met Leu Tyr Asp Pro Leu Arg Val  
 500 505 510

<sup>50</sup> Lys Pro Pro Val Glu Leu Ala Ile His Phe Ser Glu Ile Leu Glu Ile  
 515 520 525

<sup>55</sup> Pro Leu His Pro Tyr Tyr Thr Leu Tyr Trp Asn Thr Val Asn Pro Lys  
 530 535 540

Asp Val Glu Arg Leu Trp Gly Val Leu Lys Asp Lys Ala Thr Ile Glu  
 545 550 555 560

<sup>60</sup> Trp Gly Thr Phe Arg Gly Ile Lys Phe Ala Lys Lys Ile Glu Ile Ser  
 565 570 575

DE 198 40 771 A 1

Leu Asp Asp Leu Gly Ser Leu Lys Arg Thr Leu Glu Leu Leu Gly Leu				
580	585	590		
Pro His Thr Val Arg Glu Gly Ile Val Val Val Asp Tyr Pro Trp Ser				5
595	600	605		
Ala Ala Leu Leu Thr Pro Leu Gly Asn Leu Glu Trp Glu Phe Lys Ala				10
610	615	620		
Lys Pro Phe Tyr Thr Val Ile Asp Ile Ile Asn Glu Asn Asn Gin Ile				
625	630	635	640	15
Lys Leu Arg Asp Arg Gly Ile Ser Trp Ile Gly Ala Arg Met Gly Arg				
645	650	655		
Pro Glu Lys Ala Lys Glu Arg Lys Met Lys Pro Pro Val Gln Val Leu				20
660	665	670		
Phe Pro Ile Gly Leu Ala Gly Gly Ser Ser Arg Asp Ile Lys Lys Ala				25
675	680	685		
Ala Glu Glu Gly Lys Ile Ala Glu Val Glu Ile Ala Phe Phe Lys Cys				30
690	695	700		
Pro Lys Cys Gly His Val Gly Pro Glu Thr Leu Cys Pro Glu Cys Gly				35
705	710	715	720	
Ile Arg Lys Glu Leu Ile Trp Thr Cys Pro Lys Cys Gly Ala Glu Tyr				
725	730	735		
Thr Asn Ser Gln Ala Glu Gly Tyr Ser Tyr Ser Cys Pro Lys Cys Asn				40
740	745	750		
Val Lys Leu Lys Pro Phe Thr Lys Arg Lys Ile Lys Pro Ser Glu Leu				45
755	760	765		
Leu Asn Arg Ala Met Glu Asn Val Lys Val Tyr Gly Val Asp Lys Leu				
770	775	780		
Lys Gly Val Met Gly Met Thr Ser Gly Trp Lys Ile Ala Glu Pro Leu				50
785	790	795	800	
Glu Lys Gly Leu Leu Arg Ala Lys Asn Glu Val Tyr Val Phe Lys Asp				
805	810	815		
Gly Thr Ile Arg Phe Asp Ala Thr Asp Ala Pro Ile Thr His Phe Arg				55
820	825	830		

DE 198 40 771 A 1

Pro Arg Glu Ile Gly Val Ser Val Glu Lys Leu Arg Glu Leu Gly Tyr  
 835                    840                    845

5    Thr His Asp Phe Glu Gly Lys Pro Leu Val Ser Glu Asp Gln Ile Val  
 850                    855                    860

10    Glu Leu Lys Pro Gln Asp Val Ile Leu Ser Lys Glu Ala Gly Lys Tyr  
 865                    870                    875                    880

15    Leu Leu Arg Val Ala Arg Phe Val Asp Asp Leu Leu Glu Lys Phe Tyr  
 885                    890                    895

Gly Leu Pro Arg Phe Tyr Asn Ala Glu Lys Met Glu Asp Leu Ile Gly  
 900                    905                    910

20    His Leu Val Ile Gly Leu Ala Pro His Thr Ser Ala Gly Ile Val Gly  
 915                    920                    925

25    Arg Ile Ile Gly Phe Val Asp Ala Leu Val Gly Tyr Ala His Pro Tyr  
 930                    935                    940

30    Phe His Ala Ala Lys Arg Arg Asn Cys Asp Gly Asp Glu Asp Ser Val  
 945                    950                    955                    960

35    Met Leu Leu Leu Asp Ala Leu Leu Asn Phe Ser Arg Tyr Tyr Leu Pro  
 965                    970                    975

Glu Lys Arg Gly Gly Lys Met Asp Ala Pro Leu Val Ile Thr Thr Arg  
 980                    985                    990

40    Leu Asp Pro Arg Glu Val Asp Ser Glu Val His Asn Met Asp Val Val  
 995                    1000                    1005

45    Arg Tyr Tyr Pro Leu Glu Phe Tyr Glu Ala Thr Tyr Glu Leu Lys Ser  
 1010                    1015                    1020

50    Pro Lys Glu Leu Val Arg Val Ile Glu Gly Val Glu Asp Arg Leu Gly  
 1025                    1030                    1035                    1040

55    Lys Pro Glu Met Tyr Tyr Gly Ile Lys Phe Thr His Asp Thr Asp Asp  
 1045                    1050                    1055

Ile Ala Leu Gly Pro Lys Met Ser Leu Tyr Lys Gln Leu Gly Asp Met  
 1060                    1065                    1070

60    Glu Glu Lys Val Lys Arg Gln Leu Thr Leu Ala Glu Arg Ile Arg Ala  
 1075                    1080                    1085

DE 198 40 771 A 1

Val Asp Gln His Tyr Val Ala Glu Thr Ile Leu Asn Ser His Leu Ile			
1090	1095	1100	
Pro Asp Leu Arg Gly Asn Leu Arg Ser Phe Thr Arg Gln Glu Phe Arg			5
1105	1110	1115	1120
Cys Val Lys Cys Asn Thr Lys Tyr Arg Arg Pro Pro Leu Asp Gly Lys			10
1125	1130	1135	
Cys Pro Val Cys Gly Gly Lys Ile Val Leu Thr Val Ser Lys Gly Ala			15
1140	1145	1150	
Ile Glu Lys Tyr Leu Gly Thr Ala Lys Met Leu Val Ala Asn Tyr Asn			20
1155	1160	1165	
Val Lys Pro Tyr Thr Arg Gln Arg Ile Cys Leu Thr Glu Lys Asp Ile			25
1170	1175	1180	
Asp Ser Leu Phe Glu Tyr Leu Phe Pro Glu Ala Gln Leu Thr Leu Ile			30
1185	1190	1195	1200
Val Asp Pro Asn Asp Ile Cys Met Lys Met Ile Lys Glu Arg Thr Gly			35
1205	1210	1215	
Glu Thr Val Gln Gly Gly Leu Leu Glu Asn Ser Ser Gly Asn			40
1220	1225	1230	
Asn Gly Lys Lys Ile Glu Lys Lys Glu Lys Lys Ala Lys Glu Lys Pro			45
1235	1240	1245	
Lys Lys Lys Lys Val Ile Ser Leu Asp Asp Phe Phe Ser Lys Arg			50
1250	1255	1260	
(2) ANGABEN ZU SEQ ID NO: 32:			
(i) SEQUENZKENNZEICHEN:			
(A) LÄNGE: 363 Aminosäuren			
(B) ART: Aminosäure			
(C) STRANGFORM: Einzelstrang			
(D) TOPOLOGIE: linear			
(ii) ART DES MOLEKÜLS: Protein			
(vi) URSPRÜNLICHE HERKUNFT:			
(A) ORGANISMUS: Homo sapiens			

DE 198 40 771 A 1

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

5 Met Gln Ala Phe Leu Lys Gly Thr Ser Ile Ser Thr Lys Pro Pro Leu  
1 5 10 15

10 Thr Lys Asp Arg Gly Val Ala Ala Ser Ala Gly Ser Ser Gly Glu Asn  
20 25 30

15 Lys Lys Ala Lys Pro Val Pro Trp Val Glu Lys Tyr Arg Pro Lys Cys  
35 40 45

20 Val Asp Glu Val Ala Phe Gln Glu Glu Val Val Ala Val Leu Lys Lys  
50 55 60

25 Ser Leu Glu Gly Ala Asp Leu Pro Asn Leu Leu Phe Tyr Gly Pro Pro  
65 70 75 80

30 Gly Thr Gly Lys Thr Ser Thr Ile Leu Ala Ala Ala Arg Glu Leu Phe  
85 90 95

35 Gly Pro Glu Leu Phe Arg Leu Arg Val Leu Glu Leu Asn Ala Ser Asp  
100 105 110

40 Glu Arg Gly Ile Gln Val Val Arg Glu Lys Val Lys Asn Phe Ala Gln  
115 120 125

45 Leu Thr Val Ser Gly Ser Arg Ser Asp Gly Lys Pro Cys Pro Pro Phe  
130 135 140

50 Lys Ile Val Ile Leu Asp Glu Ala Asp Ser Met Thr Ser Ala Ala Gln  
145 150 155 160

55 Ala Ala Leu Arg Arg Thr Met Glu Lys Glu Ser Lys Thr Thr Arg Phe  
165 170 175

60 Cys Leu Ile Cys Asn Tyr Val Ser Arg Ile Ile Glu Pro Leu Thr Ser  
180 185 190

65 Arg Cys Ser Lys Phe Arg Phe Lys Pro Leu Ser Asp Lys Ile Gln Gln  
195 200 205

70 Gln Arg Leu Leu Asp Ile Ala Lys Lys Glu Asn Val Lys Ile Ser Asp  
210 215 220

75 Glu Gly Ile Ala Tyr Leu Val Lys Val Ser Glu Gly Asp Leu Arg Lys  
225 230 235 240

65

DE 198 40 771 A 1

Ala Ile Thr Phe Leu Gln Ser Ala Thr Arg Leu Thr Gly Gly Lys Glu				
245	250	255		
Ile Thr Glu Lys Val Ile Thr Asp Ile Ala Gly Val Ile Pro Ala Glu				5
260	265	270		
Lys Ile Asp Gly Val Phe Ala Ala Cys Gln Ser Gly Ser Phe Asp Lys				10
275	280	285		
Leu Glu Ala Val Val Lys Asp Leu Ile Asp Glu Gly His Ala Ala Thr				15
290	295	300		
Gln Leu Val Asn Gln Leu His Asp Val Val Val Glu Asn Asn Leu Ser				20
305	310	315	320	
Asp Lys Gln Lys Ser Ile Ile Thr Glu Lys Leu Ala Glu Val Asp Lys				
325	330	335		
Cys Leu Ala Asp Gly Ala Asp Glu His Leu Gln Leu Ile Ser Leu Cys				25
340	345	350		
Ala Thr Val Met Gln Gln Leu Ser Gln Asn Cys				30
355	360			
(2) ANGABEN ZU SEQ ID NO: 33:				
(i) SEQUENZKENNZEICHEN:				
(A) LÄNGE: 329 Aminosäuren				
(B) ART: Aminosäure				
(C) STRANGFORM: Einzelstrang				
(D) TOPOLOGIE: linear				
(ii) ART DES MOLEKÜLS: Protein				
(vi) URSPRÜNLICHE HERKUNFT:				
(A) ORGANISMUS: Homo sapiens				
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:				
Asn Leu Val Gln Cys Gly Asp Phe Pro His Leu Leu Val Tyr Gly Pro				55
1	5	10	15	
Ser Gly Ala Gly Lys Lys Thr Arg Ile Met Cys Ile Leu Arg Glu Leu				60
20	25	30		

## DE 198 40 771 A 1

Tyr Gly Val Gly Val Glu Lys Leu Arg Ile Glu His Gln Thr Ile Thr  
 35 40 45

5 Thr Pro Ser Lys Lys Ile Glu Ile Ser Thr Ile Ala Ser Asn Tyr  
 50 55 60

10 His Leu Glu Val Asn Pro Ser Asp Ala Gly Asn Ser Asp Arg Val Val  
 65 70 75 80

15 Ile Gln Glu Met Leu Lys Thr Val Ala Gln Ser Gln Gln Leu Glu Thr  
 85 90 95

20 Asn Ser Gln Arg Asp Phe Lys Val Val Leu Leu Thr Glu Val Asp Lys  
 100 105 110

25 Leu Thr Lys Asp Ala Gln His Ala Leu Arg Arg Thr Met Glu Lys Tyr  
 115 120 125

30 Met Ser Thr Cys Arg Leu Ile Leu Cys Cys Asn Ser Thr Ser Lys Val  
 130 135 140

35 Ile Pro Pro Ile Arg Ser Arg Cys Leu Ala Val Arg Val Pro Ala Pro  
 145 150 155 160

40 Ser Ile Glu Asp Ile Cys His Val Leu Ser Thr Val Cys Lys Lys Glu  
 165 170 175

45 Gly Leu Asn Leu Pro Ser Gln Leu Ala His Arg Leu Ala Glu Lys Ser  
 180 185 190

50 Cys Arg Asn Leu Arg Lys Ala Leu Leu Met Cys Glu Ala Cys Arg Val  
 195 200 205

55 Gln Gln Tyr Pro Phe Thr Ala Asp Gln Glu Ile Pro Glu Thr Asp Trp  
 210 215 220

60 Glu Val Tyr Leu Arg Glu Thr Ala Asn Ala Ile Val Ser Gln Gln Thr  
 225 230 235 240

65 Pro Gln Arg Leu Leu Glu Val Arg Gly Arg Leu Tyr Glu Leu Leu Thr  
 245 250 255

70 His Cys Ile Pro Pro Glu Ile Ile Met Lys Gly Leu Leu Ser Glu Leu  
 260 265 270

75 Leu His Asn Cys Asp Gly Gln Leu Lys Gly Glu Val Ala Gln Met Ala  
 275 280 285

DE 198 40 771 A 1

Ala Tyr Tyr Glu His Arg Leu Gln Leu Gly Ser Lys Ala Ile Tyr His  
 290 295 300

Leu Glu Ala Phe Val Ala Lys Phe Met Ala Leu Tyr Lys Lys Phe Ile  
 305 310 315 320

Gln Asp Gly Leu Glu Gly Met Met Phe  
 325

(2) ANGABEN ZU SEQ ID NO: 34:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 354 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

Met Glu Val Glu Ala Val Cys Gly Gly Ala Gly Glu Val Ala Gln  
 1 5 10 15

Asp Ser Asp Pro Ala Pro Ala Phe Ser Lys Ala Pro Gly Ser Ala Gly  
 20 25 30

His Tyr Glu Leu Pro Trp Val Glu Lys Tyr Arg Pro Val Lys Leu Asn  
 35 40 45

Glu Ile Val Gly Asn Glu Asp Thr Val Ser Arg Leu Glu Val Phe Ala  
 50 55 60

Arg Glu Gly Asn Val Pro Asn Ile Ile Ala Gly Pro Pro Gly Thr  
 65 70 75 80

Gly Lys Thr Thr Ser Ile Leu Cys Leu Ala Arg Ala Leu Leu Gly Pro  
 85 90 95

Ala Leu Lys Asp Ala Met Leu Glu Leu Asn Ala Ser Asn Asp Arg Gly  
 100 105 110

DE 198 40 771 A 1

Ile Asp Val Val Arg Asn Lys Ile Lys Met Phe Ala Gln Gln Lys Val  
115 120 125

5 Thr Leu Pro Lys Gly Arg His Lys Ile Ile Leu Asp Glu Ala Asp  
130 135 140

10 Ser Met Thr Asp Gly Ala Gln Gln Ala Leu Arg Arg Thr Met Glu Ile  
145 150 155 160

15 Tyr Ser Lys Thr Thr Arg Phe Ala Leu Ala Cys Asn Ala Ser Asp Lys  
165 170 175

20 Ile Ile Glu Pro Ile Gln Ser Arg Cys Ala Val Leu Arg Tyr Thr Lys  
180 185 190

25 Leu Thr Asp Ala Gln Ile Leu Thr Arg Leu Met Asn Val Ile Glu Lys  
195 200 205

30 Glu Arg Val Pro Tyr Thr Asp Asp Gly Leu Glu Ala Ile Ile Phe Thr  
210 215 220

35 Ala Gln Gly Asp Met Arg Gln Ala Leu Asn Asn Leu Gln Ser Thr Phe  
225 230 235 240

40 Ser Gly Phe Gly Phe Ile Asn Ser Glu Asn Val Phe Lys Val Cys Asp  
245 250 255

45 Glu Pro His Pro Leu Leu Val Lys Glu Met Ile Gln His Cys Val Asn  
260 265 270

50 Ala Asn Ile Asp Glu Ala Tyr Lys Ile Leu Ala His Leu Trp His Leu  
275 280 285

55 Gly Tyr Ser Pro Glu Asp Ile Ile Gly Asn Ile Phe Arg Val Cys Lys  
290 295 300

60 Thr Phe Gln Met Ala Glu Tyr Leu Lys Leu Glu Phe Ile Lys Glu Ile  
305 310 315 320

65 Gly Tyr Thr His Met Lys Ile Ala Glu Gly Val Asn Ser Leu Leu Gln  
325 330 335

70 Met Ala Gly Leu Leu Ala Arg Leu Cys Gln Lys Thr Met Ala Pro Val  
340 345 350

75 Ala Ser

## (2) ANGABEN ZU SEQ ID NO: 35:

## (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 366 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 (D) TOPOLOGIE: linear

5

10

## (ii) ART DES MOLEKÜLS: Protein

## (vi) URSPRÜNLICHE HERKUNFT:

(A) ORGANISMUS: Escherichia coli

15

20

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35:

Met Lys Phe Thr Val Glu Arg Glu His Leu Leu Lys Pro Leu Gln Gln  
 1 5 10 15

25

Val Ser Gly Pro Leu Gly Gly Arg Pro Thr Leu Pro Ile Leu Gly Asn  
 20 25 30

30

Leu Leu Leu Gln Val Ala Asp Gly Thr Leu Ser Leu Thr Gly Thr Asp  
 35 40 45

35

Leu Glu Met Glu Met Val Ala Arg Val Ala Leu Val Gln Pro His Glu  
 50 55 60

40

Pro Gly Ala Thr Thr Val Pro Ala Arg Lys Phe Phe Asp Ile Cys Arg  
 65 70 75 80

45

Gly Leu Pro Glu Gly Ala Glu Ile Ala Val Gln Leu Glu Gly Glu Arg  
 85 90 95

45

Met Leu Val Arg Ser Gly Arg Ser Arg Phe Ser Leu Ser Thr Leu Pro  
 100 105 110

50

Ala Ala Asp Phe Pro Asn Leu Asp Asp Trp Gln Ser Glu Val Glu Phe  
 115 120 125

55

Thr Leu Pro Gln Ala Thr Met Lys Arg Leu Ile Glu Ala Thr Gln Phe  
 130 135 140

Ser Met Ala His Gln Asp Val Arg Tyr Tyr Leu Asn Gly Met Leu Phe  
 145 150 155 160

60

65

DE 198 40 771 A 1

Glu Thr Glu Gly Glu Glu Leu Arg Thr Val Ala Thr Asp Gly His Arg  
165 170 175

5 Leu Ala Val Cys Sér Met Pro Ile Gly Gln Ser Leu Pro Ser His Ser  
180 185 190

10 Val Ile Val Pro Arg Lys Gly Val Ile Glu Leu Met Arg Met Leu Asp  
195 200 205

15 Gly Gly Asp Asn Pro Leu Arg Val Gln Ile Gly Ser Asn Asn Ile Arg  
210 215 220

20 Ala His Val Gly Asp Phe Ile Phe Thr Ser Lys Leu Val Asp Gly Arg  
225 230 235 240

25 Phe Pro Asp Tyr Arg Arg Val Leu Pro Lys Asn Pro Asp Lys His Leu  
245 250 255

30 Glu Ala Gly Cys Asp Leu Leu Lys Gln Ala Phe Ala Arg Ala Ala Ile  
260 265 270

35 Leu Ser Asn Glu Lys Phe Arg Gly Val Arg Leu Tyr Val Ser Glu Asn  
275 280 285

40 Gln Leu Lys Ile Thr Ala Asn Asn Pro Glu Gln Glu Ala Glu Glu  
290 295 300

45 Ile Leu Asp Val Thr Tyr Ser Gly Ala Glu Met Glu Ile Gly Phe Asn  
305 310 315 320

50 Val Ser Tyr Val Leu Asp Val Leu Asn Ala Leu Lys Cys Glu Asn Val  
325 330 335

55 Arg Met Met Leu Thr Asp Ser Val Ser Ser Val Gln Ile Glu Asp Ala  
340 345 350

60 Ala Ser Gln Ser Ala Ala Tyr Val Val Met Pro Met Arg Leu  
355 360 365

(2) ANGABEN ZU SEQ ID NO: 36:

55 (ii) SEQUENZKENNZEICHEN:

(A) LÄNGE: 363 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

65 (ii) ART DES MOLEKÜLS: Protein

(vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: *Aquifex Aeolicus*

5

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36:

10

Met Arg Val Lys Val Asp Arg Glu Glu Leu Glu Glu Val Leu Lys Lys  
 1 5 10 15

15

Ala Arg Glu Ser Thr Glu Lys Lys Ala Ala Leu Pro Ile Leu Ala Asn  
 20 25 30

20

Phe Leu Leu Ser Ala Lys Glu Glu Asn Leu Ile Val Arg Ala Thr Asp  
 35 40 45

25

Leu Glu Asn Tyr Leu Val Val Ser Val Lys Gly Glu Val Glu Glu Glu  
 50 55 60

25

Gly Glu Val Cys Val His Ser Gln Lys Leu Tyr Asp Ile Val Lys Asn  
 65 70 75 80

30

Leu Asn Ser Ala Tyr Val Tyr Leu His Thr Glu Gly Glu Lys Leu Val  
 85 90 95

30

Ile Thr Gly Gly Lys Ser Thr Tyr Lys Leu Pro Thr Ala Pro Ala Glu  
 100 105 110

35

Asp Phe Pro Glu Phe Pro Glu Ile Val Glu Gly Gly Glu Thr Leu Ser  
 115 120 125

40

Gly Asn Leu Leu Val Asn Gly Ile Glu Lys Val Glu Tyr Ala Ile Ala  
 130 135 140

45

Lys Glu Glu Ala Asn Ile Ala Leu Gln Gly Met Tyr Leu Arg Gly Tyr  
 145 150 155 160

50

Glu Asp Arg Ile His Phe Val Gly Ser Asp Gly His Arg Leu Ala Leu  
 165 170 175

55

Tyr Glu Pro Leu Gly Glu Phe Ser Lys Glu Leu Leu Ile Pro Arg Lys  
 180 185 190

55

Ser Leu Lys Val Leu Lys Lys Leu Ile Thr Gly Ile Glu Asp Val Asn  
 195 200 205

60

Ile Glu Lys Ser Glu Asp Glu Ser Phe Ala Tyr Phe Ser Thr Pro Glu  
 210 215 220

65

Trp Lys Leu Ala Val Arg Leu Leu Glu Gly Glu Phe Pro Asp Tyr Met  
 225 230 235 240

5 Ser Val Ile Pro Glu Glu Phe Ser Ala Glu Val Leu Phe Glu Thr Glu  
 245 250 255

10 Glu Val Leu Lys Val Leu Lys Arg Leu Lys Ala Leu Ser Glu Gly Lys  
 260 265 270

15 Val Phe Pro Val Lys Ile Thr Leu Ser Glu Asn Leu Ala Ile Phe Glu  
 275 280 285

20 Phe Ala Asp Pro Glu Phe Gly Glu Ala Arg Glu Glu Ile Glu Val Glu  
 290 295 300

25 Tyr Thr Gly Glu Pro Phe Glu Ile Gly Phe Asn Gly Lys Tyr Leu Met  
 305 310 315 320

30 Glu Ala Leu Asp Ala Tyr Asp Ser Glu Arg Val Trp Phe Lys Phe Thr  
 325 330 335

35 Thr Pro Asp Thr Ala Thr Leu Leu Glu Ala Glu Asp Tyr Glu Lys Glu  
 340 345 350

40 Pro Tyr Lys Cys Ile Ile Met Pro Met Arg Val  
 355 360

## (2) ANGABEN ZU SEQ ID NO: 37:

40 (i) SEQUENZKENNZEICHEN:  
 (A) LÄNGE: 1160 Aminosäuren  
 (B) ART: Aminosäure  
 (C) STRANGFORM: Einzelstrang  
 45 (D) TOPOLOGIE: linear

50 (ii) ART DES MOLEKÜLS: Protein

55 (vi) URSPRÜNLICHE HERKUNFT:  
 (A) ORGANISMUS: Escherichia coli

60 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37:

Met Ser Glu Pro Arg Phe Val His Leu Arg Val His Ser Asp Tyr Ser  
 1 5 10 15

## DE 198 40 771 A 1

Met Ile Asp Gly Leu Ala Lys Thr Ala Pro Leu Val Lys Lys Ala Ala				
20	25	30		5
Ala Leu Gly Met Pro Ala Leu Ala Ile Thr Asp Phe Thr Asn Leu Cys				
35	40	45		
Gly Leu Val Lys Phe Tyr Gly Ala Gly His Gly Ala Gly Ile Lys Pro				10
50	55	60		
Ile Val Gly Ala Asp Phe Asn Val Gln Cys Asp Leu Leu Gly Asp Glu				
65	70	75	80	15
Leu Thr His Leu Thr Val Leu Ala Ala Asn Asn Thr Gly Tyr Gln Asn				
85	90	95		20
Leu Thr Leu Leu Ile Ser Lys Ala Tyr Gln Arg Gly Tyr Gly Ala Ala				
100	105	110		
Gly Pro Ile Ile Asp Arg Asp Trp Leu Ile Glu Leu Asn Glu Gly Leu				25
115	120	125		
Ile Leu Leu Ser Gly Gly Arg Met Gly Asp Val Gly Arg Ser Leu Leu				
130	135	140		30
Arg Gly Asn Ser Ala Leu Val Asp Glu Cys Val Ala Phe Tyr Glu Glu				
145	150	155	160	35
His Phe Pro Asp Arg Tyr Phe Leu Glu Leu Ile Arg Thr Gly Arg Pro				
165	170	175		40
Asp Glu Glu Ser Tyr Leu His Ala Ala Val Glu Leu Ala Glu Ala Arg				
180	185	190		
Gly Leu Pro Val Val Ala Thr Asn Asp Val Arg Phe Ile Asp Ser Ser				45
195	200	205		
Asp Phe Asp Ala His Glu Ile Arg Val Ala Ile His Asp Gly Phe Thr				
210	215	220		50
Leu Asp Asp Pro Lys Arg Pro Arg Asn Tyr Ser Pro Gln Gln Tyr Met				
225	230	235	240	55
Arg Ser Glu Glu Glu Met Cys Glu Leu Phe Ala Asp Ile Pro Glu Ala				
245	250	255		60
Leu Ala Asn Thr Val Glu Ile Ala Lys Arg Cys Asn Val Thr Val Arg				
260	265	270		

## DE 198 40 771 A 1

Leu Gly Glu Tyr Phe Leu Pro Gln Phe Pro Thr Gly Asp Met Ser Thr  
 275 280 285  
 5 Glu Asp Tyr Leu Val Lys Arg Ala Lys Glu Gly Leu Glu Glu Arg Leu  
 290 295 300  
 10 Ala Phe Leu Phe Pro Asp Glu Glu Arg Leu Lys Arg Arg Pro Glu  
 305 310 315 320  
 15 Tyr Asp Glu Arg Leu Glu Thr Glu Leu Gln Val Ile Asn Gln Met Gly  
 325 330 335  
 20 Phe Pro Gly Tyr Phe Leu Ile Val Met Glu Phe Ile Gln Trp Ser Lys  
 340 345 350  
 25 Asp Asn Gly Val Pro Val Gly Pro Gly Arg Gly Ser Gly Ala Gly Ser  
 355 360 365  
 30 Leu Val Ala Tyr Ala Leu Lys Ile Thr Asp Leu Asp Pro Leu Glu Phe  
 370 375 380  
 35 Asp Leu Leu Phe Glu Arg Phe Leu Asn Pro Glu Arg Val Ser Met Pro  
 385 390 395 400  
 40 Asp Phe Asp Val Asp Phe Cys Met Glu Lys Arg Asp Gln Val Ile Glu  
 405 410 415  
 His Val Ala Asp Met Tyr Gly Arg Asp Ala Val Ser Gln Ile Ile Thr  
 420 425 430  
 45 Phe Gly Thr Met Ala Ala Lys Ala Val Ile Arg Asp Val Gly Arg Val  
 435 440 445  
 50 Leu Gly His Pro Tyr Gly Phe Val Asp Arg Ile Ser Lys Leu Ile Pro  
 450 455 460  
 55 Pro Asp Pro Gly Met Thr Leu Ala Lys Ala Phe Glu Ala Glu Pro Gln  
 465 470 475 480  
 Leu Pro Glu Ile Tyr Glu Ala Asp Glu Glu Val Lys Ala Leu Ile Asp  
 485 490 495  
 Met Ala Arg Lys Leu Glu Gly Val Thr Arg Asn Ala Gly Lys His Ala  
 500 505 510  
 60 Gly Gly Val Val Ile Ala Pro Thr Lys Ile Thr Asp Phe Ala Pro Leu  
 515 520 525

## DE 198 40 771 A 1

Tyr Cys Asp Glu Glu Gly Lys His Pro Val Thr Gln Phe Asp Lys Ser				
530	535	540		
Asp Val Glu Tyr Ala Gly Leu Val Lys Phe Asp Phe Leu Gly Leu Arg				5
545	550	555	560	
Thr Leu Thr Ile Ile Asn Trp Ala Leu Glu Met Ile Asn Lys Arg Arg				10
565	570	575		
Ala Lys Asn Gly Glu Pro Pro Leu Asp Ile Ala Ala Ile Pro Leu Asp				15
580	585	590		
Asp Lys Lys Ser Phe Asp Met Leu Gln Arg Ser Glu Thr Thr Ala Val				20
595	600	605		
Phe Gln Leu Glu Ser Arg Gly Met Lys Asp Leu Ile Lys Arg Leu Gln				
610	615	620		
Pro Asp Cys Phe Glu Asp Met Ile Ala Leu Val Ala Leu Phe Arg Pro				25
625	630	635	640	
Gly Pro Leu Gln Ser Gly Met Val Asp Asn Phe Ile Asp Arg Lys His				30
645	650	655		
Gly Arg Glu Glu Ile Ser Tyr Pro Asp Val Gln Trp Gln His Glu Ser				35
660	665	670		
Leu Lys Pro Val Leu Glu Pro Thr Tyr Gly Ile Ile Leu Tyr Gln Glu				
675	680	685		
Gln Val Met Gln Ile Ala Gln Val Leu Ser Gly Tyr Thr Leu Gly Gly				40
690	695	700		
Ala Asp Met Leu Arg Arg Ala Met Gly Lys Lys Pro Glu Glu Met				45
705	710	715	720	
Ala Lys Gln Arg Ser Val Phe Ala Glu Gly Ala Glu Lys Asn Gly Ile				50
725	730	735		
Asn Ala Glu Leu Ala Met Lys Ile Phe Asp Leu Val Glu Lys Phe Ala				
740	745	750		
Gly Tyr Gly Phe Asn Lys Ser His Ser Ala Ala Tyr Ala Leu Val Ser				55
755	760	765		
Tyr Gln Thr Leu Trp Leu Lys Ala His Tyr Pro Ala Glu Phe Met Ala				60
770	775	780		

65

DE 198 40 771 A 1

Ala Val Met Thr Ala Asp Met Asp Asn Thr Glu Lys Val Val Gly Leu  
785 790 795 800

5 Val Asp Glu Cys Trp Arg Met Gly Leu Lys Ile Leu Pro Pro Asp Ile  
805 810 815

10 Asn Ser Gly Leu Tyr His Phe His Val Asn Asp Asp Gly Glu Ile Val  
820 825 830

15 Tyr Gly Ile Gly Ala Ile Lys Gly Val Gly Glu Gly Pro Ile Glu Ala  
835 840 845

20 Ile Ile Glu Ala Arg Asn Lys Gly Gly Tyr Phe Arg Glu Leu Phe Asp  
850 855 860

25 Leu Cys Ala Arg Thr Asp Thr Lys Lys Leu Asn Arg Arg Val Leu Glu  
865 870 875 880

30 Lys Leu Ile Met Ser Gly Ala Phe Asp Arg Leu Gly Pro His Arg Ala  
885 890 895

35 Ala Leu Met Asn Ser Leu Gly Asp Ala Leu Lys Ala Ala Asp Gln His  
900 905 910

40 Ala Lys Ala Glu Ala Ile Gly Gln Ala Asp Met Phe Gly Val Leu Ala  
915 920 925

45 Glu Glu Pro Glu Gln Ile Glu Gln Ser Tyr Ala Ser Cys Gln Pro Trp  
930 935 940

50 Pro Glu Gln Val Val Leu Asp Gly Glu Arg Glu Thr Leu Gly Leu Tyr  
945 950 955 960

55 Leu Thr Gly His Pro Ile Asn Gln Tyr Leu Lys Glu Ile Glu Arg Tyr  
965 970 975

60 Val Gly Gly Val Arg Leu Lys Asp Met His Pro Thr Glu Arg Gly Lys  
980 985 990

65 Val Ile Thr Ala Ala Gly Leu Val Val Ala Ala Arg Val Met Val Thr  
995 1000 1005

70 Lys Arg Gly Asn Arg Ile Gly Ile Cys Thr Leu Asp Asp Arg Ser Gly  
1010 1015 1020

75 Arg Leu Glu Val Met Leu Phe Thr Asp Ala Leu Asp Lys Tyr Gln Gln  
1025 1030 1035 1040

DE 198 40 771 A 1

Leu Leu Glu Lys Asp Arg Ile Leu Ile Val Ser Gly Gln Val Ser Phe					
1045	1050	1055			
Asp Asp Phe Ser Gly Gly Leu Lys Met Thr Ala Arg Glu Val Met Asp					5
1060	1065	1070			
Ile Asp Glu Ala Arg Glu Lys Tyr Ala Arg Gly Leu Ala Ile Ser Leu					10
1075	1080	1085			
Thr Asp Arg Gln Ile Asp Asp Gln Leu Leu Asn Arg Leu Arg Gln Ser					15
1090	1095	1100			
Leu Glu Pro His Arg Ser Gly Thr Ile Pro Val His Leu Tyr Tyr Gln					
1105	1110	1115	1120		
Arg Ala Asp Ala Arg Ala Arg Leu Arg Phe Gly Ala Thr Trp Arg Val					20
1125	1130	1135			
Ser Pro Ser Asp Arg Leu Leu Asn Asp Leu Arg Gly Leu Ile Gly Ser					25
1140	1145	1150			
Glu Gln Val Glu Leu Glu Phe Asp					30
1155	1160				
(2) ANGABEN ZU SEQ ID NO: 38:					
(i) SEQUENZKENNZEICHEN:					
(A) LÄNGE: 1161 Aminosäuren					
(B) ART: Aminosäure					
(C) STRANGFORM: Einzelstrang					
(D) TOPOLOGIE: linear					
(ii) ART DES MOLEKÜLS: Protein					
(vi) URSPRÜNLICHE HERKUNFT:					
(A) ORGANISMUS: Aquifex Aeolicus					
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 38:					
Met Ser Lys Asp Phe Val His Leu His Leu His Thr Gln Phe Ser Leu					55
1	5	10	15		
Leu Asp Gly Ala Ile Lys Ile Asp Glu Leu Val Lys Lys Ala Lys Glu					60
20	25	30			

DE 198 40 771 A 1

Tyr Gly Tyr Lys Ala Val Gly Met Ser Asp His Gly Asn Leu Phe Gly  
35 40 45

5 Ser Tyr Lys Phe Tyr Lys Ala Leu Lys Ala Glu Gly Ile Lys Pro Ile  
50 55 60

10 Ile Gly Met Glu Ala Tyr Phe Thr Thr Gly Ser Arg Phe Asp Arg Lys  
65 70 75 80

15 Thr Lys Thr Ser Glu Asp Asn Ile Thr Asp Lys Tyr Asn His His Leu  
85 90 95

20 Ile Leu Ile Ala Lys Asp Asp Lys Gly Leu Lys Asn Leu Met Lys Leu  
100 105 110

25 Ser Thr Leu Ala Tyr Lys Glu Gly Phe Tyr Tyr Lys Pro Arg Ile Asp  
115 120 125

30 Tyr Glu Leu Leu Glu Lys Tyr Gly Glu Gly Leu Ile Ala Leu Thr Ala  
130 135 140

35 Cys Leu Lys Gly Val Pro Thr Tyr Tyr Ala Ser Ile Asn Glu Val Lys  
145 150 155 160

40 Lys Ala Glu Glu Trp Val Lys Lys Phe Lys Asp Ile Phe Gly Asp Asp  
165 170 175

45 Leu Tyr Leu Glu Leu Gln Ala Asn Asn Ile Pro Glu Gln Glu Val Ala  
180 185 190

50 Asn Arg Asn Leu Ile Glu Ile Ala Lys Lys Tyr Asp Val Lys Leu Ile  
195 200 205

55 Ala Thr Gln Asp Ala His Tyr Leu Asn Pro Glu Asp Arg Tyr Ala His  
210 215 220

60 Thr Val Leu Met Ala Leu Gln Met Lys Lys Thr Ile His Glu Leu Ser  
225 230 235 240

65 Ser Gly Asn Phe Lys Cys Ser Asn Glu Asp Leu His Phe Ala Pro Pro  
245 250 255

70 Glu Tyr Met Trp Lys Lys Phe Glu Gly Lys Phe Glu Gly Trp Glu Lys  
260 265 270

75 Ala Leu Leu Asn Thr Leu Glu Val Met Glu Lys Thr Ala Asp Ser Phe  
275 280 285

## DE 198 40 771 A 1

Glu	Ile	Phe	Glu	Asn	Ser	Thr	Tyr	Leu	Leu	Pro	Lys	Tyr	Asp	Val	Pro	
290			295			300										5
Pro	Asp	Lys	Thr	Leu	Glu	Glu	Tyr	Leu	Arg	Glu	Leu	Ala	Tyr	Lys	Gly	
305			310			315			320							10
Leu	Arg	Gln	Arg	Ile	Glu	Arg	Gly	Gln	Ala	Lys	Asp	Thr	Lys	Glu	Tyr	
325			330			335										15
Trp	Glu	Arg	Leu	Glu	Tyr	Glu	Leu	Glu	Val	Ile	Asn	Lys	Met	Gly	Phe	
340			345			350										20
Ala	Gly	Tyr	Phe	Leu	Ile	Val	Gln	Asp	Phe	Ile	Asn	Trp	Ala	Lys	Lys	
355			360			365										25
Asn	Asp	Ile	Pro	Val	Gly	Pro	Gly	Arg	Gly	Ser	Ala	Gly	Gly	Ser	Leu	
370			375			380										30
Val	Ala	Tyr	Ala	Ile	Gly	Ile	Thr	Asp	Val	Asp	Pro	Ile	Lys	His	Gly	
385			390			395			400							35
Phe	Leu	Phe	Glu	Arg	Phe	Leu	Asn	Pro	Glu	Arg	Val	Ser	Met	Pro	Asp	
405			410			415										40
Ile	Asp	Val	Asp	Phe	Cys	Gln	Asp	Asn	Arg	Glu	Lys	Val	Ile	Glu	Tyr	
420			425			430										45
Val	Arg	Asn	Lys	Tyr	Gly	His	Asp	Asn	Val	Ala	Gln	Ile	Ile	Thr	Tyr	
435			440			445										50
Asn	Val	Met	Lys	Ala	Lys	Gln	Thr	Leu	Arg	Asp	Val	Ala	Arg	Ala	Met	
450			455			460										55
Gly	Leu	Pro	Tyr	Ser	Thr	Ala	Asp	Lys	Leu	Ala	Lys	Leu	Ile	Pro	Gln	
465			470			475			480							60
Gly	Asp	Val	Gln	Gly	Thr	Trp	Leu	Ser	Leu	Glu	Glu	Met	Tyr	Lys	Thr	
485			490			495										65
Pro	Val	Glu	Glu	Leu	Leu	Gln	Lys	Tyr	Gly	Glu	His	Arg	Thr	Asp	Ile	
500			505			510										70
Glu	Asp	Asn	Val	Lys	Lys	Phe	Arg	Gln	Ile	Cys	Glu	Glu	Ser	Pro	Glu	
515			520			525										75
Ile	Lys	Gln	Leu	Val	Glu	Thr	Ala	Leu	Lys	Leu	Glu	Gly	Leu	Thr	Arg	
530			535			540										80

DE 198 40 771 A 1

His Thr Ser Leu His Ala Ala Gly Val Val Ile Ala Pro Lys Pro Leu  
 545 550 555 560

<sup>5</sup> Ser Glu Leu Val Pro Leu Tyr Tyr Asp Lys Glu Gly Glu Val Ala Thr  
 565 570 575

<sup>10</sup> Gln Tyr Asp Met Val Gln Leu Glu Glu Leu Gly Leu Leu Lys Met Asp  
 580 585 590

Phe Leu Gly Leu Lys Thr Leu Thr Glu Leu Lys Leu Met Lys Glu Leu  
<sup>15</sup> 595 600 605

Ile Lys Glu Arg His Gly Val Asp Ile Asn Phe Leu Glu Leu Pro Leu  
 610 615 620

<sup>20</sup> Asp Asp Pro Lys Val Tyr Lys Leu Leu Gln Glu Gly Lys Thr Thr Gly  
 625 630 635 640

<sup>25</sup> Val Phe Gln Leu Glu Ser Arg Gly Met Lys Glu Leu Leu Lys Lys Leu  
 645 650 655

<sup>30</sup> Lys Pro Asp Ser Phe Asp Asp Ile Val Ala Val Leu Ala Leu Tyr Arg  
 660 665 670

<sup>35</sup> Pro Gly Pro Leu Lys Ser Gly Leu Val Asp Thr Tyr Ile Lys Arg Lys  
 675 680 685

His Gly Lys Glu Pro Val Glu Tyr Pro Phe Pro Glu Leu Glu Pro Val  
 690 695 700

<sup>40</sup> Leu Lys Glu Thr Tyr Gly Val Ile Val Tyr Gln Glu Gln Val Met Lys  
 705 710 715 720

<sup>45</sup> Met Ser Gln Ile Leu Ser Gly Phe Thr Pro Gly Glu Ala Asp Thr Leu  
 725 730 735

<sup>50</sup> Arg Lys Ala Ile Gly Lys Lys Ala Asp Leu Met Ala Gln Met Lys  
 740 745 750

<sup>55</sup> Asp Lys Phe Ile Gln Gly Ala Val Glu Arg Gly Tyr Pro Glu Glu Lys  
 755 760 765

Ile Arg Lys Leu Trp Glu Asp Ile Glu Lys Phe Ala Ser Tyr Ser Phe  
 770 775 780

<sup>60</sup> Asn Lys Ser His Ser Val Ala Tyr Gly Tyr Ile Ser Tyr Trp Thr Ala  
 785 790 795 800

## DE 198 40 771 A 1

Tyr Val Lys Ala His Tyr Pro Ala Glu Phe Phe Ala Val Lys Leu Thr			
805	810	815	
			5
Thr Glu Lys Asn Asp Asn Lys Phe Leu Asn Leu Ile Lys Asp Ala Lys			
820	825	830	
			10
Leu Phe Gly Phe Glu Ile Leu Pro Pro Asp Ile Asn Lys Ser Asp Val			
835	840	845	
			15
Gly Phe Thr Ile Glu Gly Glu Asn Arg Ile Arg Phe Gly Leu Ala Arg			
850	855	860	
			20
Ile Lys Gly Val Gly Glu Glu Thr Ala Lys Ile Ile Val Glu Ala Arg			
865	870	875	880
			25
Lys Lys Tyr Lys Gln Phe Lys Gly Leu Ala Asp Phe Ile Asn Lys Thr			
885	890	895	
			30
Lys Asn Arg Lys Ile Asn Lys Lys Val Val Glu Ala Leu Val Lys Ala			
900	905	910	
			35
Gly Ala Phe Asp Phe Thr Lys Lys Arg Lys Glu Leu Leu Ala Lys			
915	920	925	
			40
Val Ala Asn Ser Glu Lys Ala Leu Met Ala Thr Gln Asn Ser Leu Phe			
930	935	940	
			45
Gly Ala Pro Lys Glu Glu Val Glu Leu Asp Pro Leu Lys Leu Glu			
945	950	955	960
			50
Lys Glu Val Leu Gly Phe Tyr Ile Ser Gly His Pro Leu Asp Asn Tyr			
965	970	975	
			55
Glu Lys Leu Leu Lys Asn Arg Tyr Thr Pro Ile Glu Asp Leu Glu Glu			
980	985	990	
			60
Trp Asp Lys Glu Ser Glu Ala Val Leu Thr Gly Val Ile Thr Glu Leu			
995	1000	1005	
			65
Lys Val Lys Lys Thr Lys Asn Gly Asp Tyr Met Ala Val Phe Asn Leu			
1010	1015	1020	
			70
Val Asp Lys Thr Gly Leu Ile Glu Cys Val Val Phe Pro Gly Val Tyr			
1025	1030	1035	1040
			75
Glu Glu Ala Lys Glu Leu Ile Glu Glu Asp Arg Val Val Val Lys			
1045	1050	1055	

Gly Phe Leu Asp Glu Asp Leu Glu Thr Glu Asn Val Lys Phe Val Val  
 1060 1065 1070

5 Lys Glu Val Phe Ser Pro Glu Glu Phe Ala Lys Glu Met Arg Asn Thr  
 1075 1080 1085

10 Leu Tyr Ile Phe Leu Lys Arg Glu Gln Ala Leu Asn Gly Val Ala Glu  
 1090 1095 1100

15 Lys Leu Lys Gly Ile Ile Glu Asn Asn Arg Thr Glu Asp Gly Tyr Asn  
 1105 1110 1115 1120

Leu Val Leu Thr Val Asp Leu Gly Asp Tyr Phe Val Asp Leu Ala Leu  
 1125 1130 1135

20 Pro Gln Asp Met Lys Leu Lys Ala Asp Arg Lys Val Val Glu Glu Ile  
 1140 1145 1150

25 Glu Lys Leu Gly Val Lys Val Ile Ile  
 1155 1160

30 Patentansprüche

1. Thermostabiler in vitro Komplex zur Template-abhängigen Elongation von Nukleinsäuren, umfassend ein thermostabiles Gleitklammerprotein, welches mit einem thermostabilen Polymeraseaktivität-aufweisenden Elongationsprotein verbunden ist.
2. Thermostabiler Komplex nach Anspruch 1, dadurch gekennzeichnet, dass das Gleitklammerprotein und das Elongationsprotein über ein Kopplungsprotein verbunden sind.
3. Thermostabiler Komplex nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass das Gleitklammerprotein und das Elongationsprotein aus Archaeabakterien stammen.
4. Thermostabiler Komplex nach Anspruch 1, 2 oder 3, dadurch gekennzeichnet, dass das Gleitklammerprotein eine ringförmige Struktur aufweist, welche die Template-Nukleinsäurestränge ganz oder teilweise umschließt.
5. Thermostabiler Komplex nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass das Gleitklammerprotein eine oder beide der folgenden Aminosäurekonsensussequenzen aufweist:  
 SEQ ID NO: 39  
 [G A V L I M P F W]-D-X-X-[G A V L I M P F W]-X-X-[G A V L I M P F W]-X-[G A V L I M P F W]-X-[G A V L I M P F W]-D und/oder  
 SEQ ID NO: 40  
 [G A V L I M P F W]-X(3)-L-A-P-[K R H D E]-[G A V L I M P F W]-E.
6. Thermostabiler Komplex nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass das Gleitklammerprotein zu den menschlichen (Eukaryonten) PCNA Aminosäuresequenz (SEQ ID NO: 11) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist und/oder das Gleitklammerprotein zu der bakterielle  $\beta$ -clamp Sequenz aus E.coli (Eubakteria) (SEQ ID NO: 35) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist und/oder das Gleitklammerprotein zu der Aminosäuresequenz des PCNA Homologen aus Archaeoglobus Fulgidus (Archaea) (SEQ ID NO: 12) auf einer Länge von mindestens 100 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist.
7. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Gleitklammerprotein mit einem aus einem Alignment aus Abb. 12 generierten Hidden Markow Modell einen Score von mindestens 20 ergibt und/oder wobei das Gleitklammerprotein mit einem aus einem Alignment aus Abb. 13 generierten Hidden Markow Modell einen Score von mindestens 25 ergibt.
8. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Gleitklammerprotein ausgewählt ist aus AF0335 aus Archaeoglobus Fulgidus, MJ0247 aus Methanococcus jannaschii, PHLA008 aus Pyrococcus horikoshii, MTH1312 aus Methanobacterium thermoautotrophicus sowie AE000761\_7 aus Aquifex aeolicus.
9. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Elongationsprotein eine 5'-3'-Polymeraseaktivität oder eine Reverse-Transkriptaseaktivität aufweist.
10. Thermostabiler Komplex nach Anspruch 9, dadurch gekennzeichnet, dass das Elongationsprotein mindestens eine der folgenden Konsensussequenzen beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht:

SEQ ID NO: 44  
D-[G A V L I M P F W]-[G A V L I M P F W]-X-X-Y-N-X-X-F-D-X-P-Y-[G A V K U N O F W]-X-X-R-A

SEQ ID NO: 45  
SEQ ID NO: 46  
A-[G A V L I M P F W]-R-I-A-[F A V L I M P F W]-A-[G A V L I M P F W]-[G A V L I M P F W]-I-E-G-[G A V L I M P F W]-V-X-A-P-[G A V L I M P F W]-E-G-I-A-X-V-[K R II D E]-I

[G A V L I M P F W]-P-V-G-[G A V L I M P F W]-G-R-G-S-X-[G A V L I M P F W]-G-S-[G A V K U N O F W]-V-A-X-A-[G A V L I M P F W]-X-I-T-D-[G A V K U N O F W]-D-P-[G A V L I M P F W]-X-X-X-[G A V L I M P F W]-L-F-E-R-F-L-N-P-E-R-[G A V L I M P F W]-S-M-P-D.

11. Thermostabiler Komplex nach Anspruch 9, dadurch gekennzeichnet, dass das Elongationsprotein eine zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 22) auf einer Länge von mindestens 200 Aminosäuren bei einem Sequenzalignment eine mindestens 20%ige Sequenzidentität aufweist und/oder zur archaebakteriellen Aminosäuresequenz (SEQ ID NO: 27) auf einer Länge von mindestens 400 Aminosäuren bei einem Sequenzalignment eine mindestens 25%ige Sequenzidentität aufweist und/oder zur eubakteriellen Aminosäuresequenz (SEQ ID NO: 37) auf einer Länge von mindestens 300 Aminosäuren bei einem Sequenzalignment eine mindestens 25%ige Sequenzidentität aufweist. 10

12. Thermostabiler Komplex nach Anspruch 9, dadurch gekennzeichnet, dass das Elongationsprotein mit einem aus einem Alignment aus Abb. 17 generierten Hidden Markow Modell einen Score von mindestens 20 ergibt und/oder wobei das Elongationsprotein mit einem aus einem Alignment aus Abb. 18 generierten Hidden Markow Modell einen Score von mindestens 35 ergibt und/oder wobei das Elongationsprotein mit einem aus einem Alignment aus Abb. 19 generierten Hidden Markow Modell einen Score von mindestens 20 ergibt. 20

13. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Elongationsprotein ausgewählt ist aus AF0497 oder AF1722 aus *Archaeoglobus fulgidus*, MJ0885 oder MJ1630 aus *Methanococcus jannaschii*, PHBT047 oder PHBN021 aus *Pyrococcus horikoschii*, MTH1208 oder MTH1536 aus *Methanobacterium thermoautotrophicum* sowie PFUORF3 aus *Pyrococcus furiosus*. 25

14. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Kopplungsprotein die folgende Konsensussequenz beinhaltet und an nicht mehr als vier Positionen von dieser Sequenz abweicht:

(SEQ ID NO: 43)  
[FL]-[G A V L I M P F W]-X-X-[G A V L I M P F W]-X-G-X(13)-[G A V L I M P F W]-X-[YR]-[G A V L I M P F W]-X-[G A V L I M P F W]-A-G-[DN]-[G A V L I M P F W]-[G A V L I M P F W]-[DS].

15. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Kopplungsprotein zu der menschlichen (Eukaryonten) Aminosäuresequenz (SEQ ID NO: 16) auf einer Länge von mindestens 150 Aminosäuren bei einem Sequenzalignment eine mindestens 18%ige Sequenzidentität aufweist. 30

16. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Kopplungsprotein mit einem aus einem Alignment aus Abb. 16 generierten Hidden Markow Modell einen Score von mindestens 10 ergibt. 35

17. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass das Kopplungsprotein ausgewählt ist aus AF1790 aus *Archaeoglobus fulgidus*, MJ0702 aus *Methanococcus jannaschii*, PHBN023 aus *Pyrococcus horikoschii*, MTH1405 aus *Methanobacterium thermoautotrophicum* sowie PFUORF2 aus *Pyrococcus furiosus*. 40

18. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass der Komplex mit einem Protein assoziiert ist, welches als Gleitklammerlader fungiert. 45

19. Thermostabiler Komplex nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass der Komplex mit ATP oder einem anderen Kofaktor assoziiert vorliegt.

20. Thermostabiler akzessorischer in vitro-Komplex, dadurch gekennzeichnet, dass er ein Gleitklammerprotein und ein Kopplungsprotein enthält, wie sie jeweils in einem der vorhergehenden Ansprüche definiert sind. 50

21. Rekombinante DNA-Sequenz, dadurch gekennzeichnet, dass sie für einen thermostabilen in vitro-Komplex gemäß einem der Ansprüche 1 bis 19 oder für einen thermostabilen akzessorischen Komplex nach Anspruch 20 codiert.

22. Vektor, dadurch gekennzeichnet, dass er eine für ein Gleitklammerprotein und ein Kopplungsprotein oder/und ein Elongationsprotein codierende, rekombinante DNA-Sequenz enthält. 55

23. Vektor nach Anspruch 22, dadurch gekennzeichnet, dass er zusätzlich geeignete Restriktionsschnittstellen zur Insertion weiterer DNA-Sequenzen in einer solchen Anordnung enthält, dass sich ein Fusionsprotein aus Gleitklammerprotein und dem Expressionsprodukt der weiteren DNA-Sequenzen ergibt.

24. Vektor nach Anspruch 22 oder 23, dadurch gekennzeichnet, dass er geeignete, die Expression der DNA-Sequenz(en) steuernde Promotor- oder/und Operatorbereiche enthält. 60

25. Vektor nach Anspruch 24, dadurch gekennzeichnet, dass er mehrere Promotor- oder/und Operatorbereiche zur getrennten Expression mehrerer DNA-Sequenzen enthält.

26. Vektor nach einem der Ansprüche 22 bis 25, dadurch gekennzeichnet, dass er reprimier- und induzierbare Promotor- oder/und Operatorbereiche enthält.

27. Vektor nach einem der Ansprüche 22 bis 26, dadurch gekennzeichnet, dass er eine DNA-Sequenz nach Anspruch 21 enthält.

28. Wirtszelle, dadurch gekennzeichnet, dass sie mit einem oder mehreren Vektoren nach einem der Ansprüche 22 bis 27 transformiert ist. 65

29. Verfahren zur Herstellung eines thermostabilen in vitro-Komplexes gemäß einem der Ansprüche 1 bis 19 oder eines thermostabilen, akzessorischen in vitro-Komplexes gemäß Anspruch 20, dadurch gekennzeichnet, dass man eine entsprechende rekombinante DNA-Sequenz gemäß Anspruch 21 oder einen oder mehrere entsprechende Vek-

toren gemäß einem der Ansprüche 22 bis 27 in eine Wirtszelle einbringt, die Expression der Proteine bewirkt und diese aus dem Kulturmedium oder nach Zellaufschluss isoliert und gegebenenfalls noch mit weiteren Komplexbestandteilen koppelt.

5 30. Verwendung eines thermostabilen, akzessorischen in vitro-Komplexes nach Anspruch 20 zur Herstellung eines thermostabilen in vitro-Komplexes nach einem der Ansprüche 1 bis 19, dadurch gekennzeichnet, dass man den akzessorischen Komplex mit einem Polymeraseaktivität aufweisendem Elongationsprotein koppelt.

10 31. Verfahren zur Template-abhängigen Elongation von Nukleinsäuren, wobei die Nukleinsäure nötigenfalls denaturiert, mit mindestens einem Primer unter Hybridisierungsbedingungen verschmilzt, wobei der Primer komplementär zu einem flankierenden Bereich einer gewünschten Nukleinsäuresequenz des Template-Strangs ist und mit Hilfe einer Polymerase in Anwesenheit von Nukleotiden eine Primer-Elongation erfolgt, dadurch gekennzeichnet, dass als Polymerase ein thermostabiler in vitro-Komplex gemäß einem der Ansprüche 1 bis 19 verwendet wird.

15 32. Verfahren nach Anspruch 31, dadurch gekennzeichnet, dass man zur Amplifikation von DNA-Sequenzen zwei die gewünschte Nukleinsäuresequenz flankierende Primer und Deoxynukleotide oder/und Derivate davon oder/und Ribonukleotide oder/und Derivate davon verwendet.

20 33. Verfahren nach Anspruch 32, dadurch gekennzeichnet, dass man eine Polymerase-Ketten-Reaktion durchführt.

34. Verfahren nach Anspruch 31, dadurch gekennzeichnet, dass man zur Reversen-Transkription von RNA in DNA einen thermostabilen in vitro-Komplex gemäß einem der Ansprüche 1 bis 19 einsetzt, dessen Elongationsprotein Reverse-Transkriptase-Aktivität aufweist.

25 35. Verfahren nach einem der Ansprüche 31 bis 34, dadurch gekennzeichnet, dass man zur Sequenzierung von Nukleinsäuren ausgehend von einem Primer, der zu einem der zu sequenziierenden Nukleinsäure benachbarten Bereich komplementär ist, eine Template-abhängige Elongation bzw. Reverse-Transkription unter Verwendung von Deoxynukleotiden und Dideoxynukleotiden oder deren jeweiligen Derivaten gemäß der Methode von Sanger durchführt.

30 36. Verfahren nach einem der Ansprüche 31 bis 35, dadurch gekennzeichnet, dass man bei der Elongation der Nukleinsäuren Markierungen einfügt.

37. Verfahren nach Anspruch 36, dadurch gekennzeichnet, dass man markierte Primer oder/und markierte Deoxynukleotide oder/und deren Derivate oder/und markierte Dideoxynukleotide oder/und deren Derivate oder/und markierte Ribonukleotide oder/und deren Derivate einsetzt.

38. Verfahren zur Markierung von Nukleinsäuren durch Erzeugung einzelner Brüche in Phosphodiesterbindungen der Nukleinsäurekette und Ersatz eines Nukleotids an den Bruchstellen durch ein markiertes Nukleotid mit Hilfe einer Polymerase, dadurch gekennzeichnet, dass als Polymerase ein thermostabiler in vitro-Komplex nach einem der Ansprüche 1 bis 19 eingesetzt wird.

39. Reagenzien-Kit zur Elongation oder/und Amplifikation oder/und Reversen-Transkription oder/und Sequenzierung oder/und Markierung von Nukleinsäuren, enthaltend in einem oder mehreren getrennten Behältern

40 a) einen thermostabilen in vitro-Komplex gemäß einem der Ansprüche 1 bis 19 oder

41 b) einen thermostabilen akzessorischen in vitro-Komplex gemäß Anspruch 20 und gegebenenfalls separat davon ein Polymeraseaktivität aufweisendes Elongationsprotein,

42 sowie gegebenenfalls Primer, Puffersubstanzen, Nukleotide, ATP, einen oder mehrere andere Kofaktoren oder/und Pyrophosphat.

43. Kit nach Anspruch 39, dadurch gekennzeichnet, dass er zur Amplifikation von Nukleinsäuren neben den Substanzen a) oder b), welche 5'-3'-Polymeraseaktivität aufweisen, Deoxynukleotide oder/und Derivate davon enthält.

44. Kit nach Anspruch 39, dadurch gekennzeichnet, dass er zur Reversen Transkription Substanzen a) oder b) enthält, welche Reverse Transkriptase-Aktivität aufweisen, sowie Deoxynukleotide oder/und Derivate davon.

45. Kit nach einem der Ansprüche 39 bis 41, dadurch gekennzeichnet, dass er zur Sequenzierung neben Deoxynukleotiden oder Ribonukleotiden oder/und deren Derivaten, Dideoxynukleotide oder/und deren Derivate enthält.

Hierzu 15 Seite(n) Zeichnungen

50

55

60

65

**- Leerseite -**

## Abbildung 1:

Komponenten des Elongations-Enzymes	Eukaryota Homo sapiens	Archaeon Archaeoglobus Fulgidus (%Identität   Alignmentlänge) <sup>1</sup>	Archaeon Methanococcus Jannaschii (%Identität)	Archaeon Pyrococcus Horikoshii (%Identität) <sup>2</sup>	Archaeon Pyrococcus Furiosus (%Identität) <sup>1</sup>
Glicithiamin-Index 1	AC11_HUMAN [SEQ ID NO:1] AC12_HUMAN [SEQ ID NO:32] AC13_HUMAN [SEQ ID NO:33] AC14_HUMAN [SEQ ID NO:34]	AF2060 (39%   307) [SEQ ID NO:2] AF2060 (45%) [SEQ ID NO:2] AF2060 (27%) [SEQ ID NO:2] AF2060 (42%) [SEQ ID NO:2]	MJ1422 (32%   173) [SEQ ID NO:3]	PHBN012 (39%   269) [SEQ ID NO:4]	MTH0241 (35%   317) [SEQ ID NO:5]
Glicithiamin-Index 11	AC15_HUMAN [SEQ ID NO:6]	AF1195 (29%   276) [SEQ ID NO:7]	MJ0884 (25%   571) [SEQ ID NO:8]	PHBN013 (21%   429) [SEQ ID NO:9]	MTH0240 (26%   308) [SEQ ID NO:10]
Glicithiamin	PCNA_HUMAN [SEQ ID NO:11]	AF0335 (24%   257) [SEQ ID NO:12]	MJ10247 (25%   256) [SEQ ID NO:13]	PHLA008 (25%   245) [SEQ ID NO:14]	MTH1312 (28%   260) [SEQ ID NO:15]
Koppende Unterheit	DPD2_HUMAN [SEQ ID NO:16]	AF1790 (23%   211) [SEQ ID NO:17]	MJ0702 (19%   140) [SEQ ID NO:18]	PHBN023 (21%   401) <sup>3</sup> [SEQ ID NO:19]	MTH1405 (21%   265) [SEQ ID NO:20]
Elongationsprot.1 (Polymersc)	DPOD_HUMAN [SEQ ID NO:22]	AF0497 (25%   742) [SEQ ID NO:23]	MJ0885 (28%   277) [SEQ ID NO:24]	PHBT047 (29%   439) [SEQ ID NO:25]	MTH1208 (26%   650) [SEQ ID NO:26]
Elongationsprot.1 (Pol.Aktivit)	-----	AF1722 -----	MJ1630 (50%   1160) <sup>2</sup> [SEQ ID NO:27]	PHBN021 (47%   974) <sup>2</sup> [SEQ ID NO:28]	MTH1536 (49%   1139) <sup>2</sup> [SEQ ID NO:30]

## Abbildung 2:

Komponenten des Elongations-holoenzymes	Eubakteria Escherichia coli	Aquifex Aeolicus (%Identität   Alignmentlänge) <sup>1</sup>
Gleitklammer	DP3B_ECOLI [SEQ ID NO:35]	AASEQ93 (28%   370) [SEQ ID NO:36]
Elongationsprotein	DP3A_ECOLI [SEQ ID NO:37]	AASEQ50 (37%   1154) [SEQ ID NO:38]

Abbildung 3:

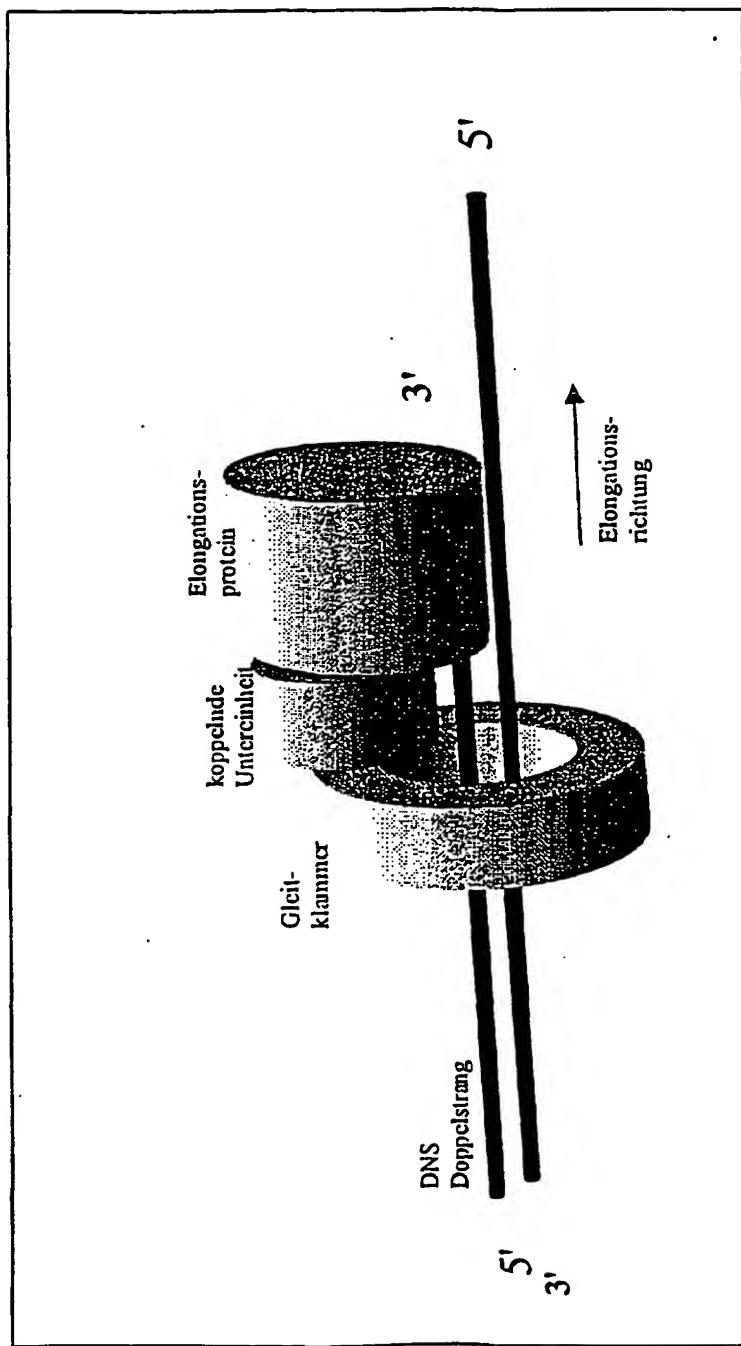


Abbildung 4:

## Gleitklammer Region 1:

PCNA\_HUMAN : MDSSHVSLVQLTLRSEGFDTYRCD  
PCNA\_METJA : MDPSHVALVSLEIPRLAEEYEAD  
MTH1312 : LDRSHITYVHLELKAELFDEYVCD  
PHLA008 : MDPSRVVLIDLNLPSIIFSKYEVD  
AF0335 : VDPANVAMIVDIPKDSFEVYNID

Konsensus : \$DXXX\$XX\$X\$X\$XXXXFXXYXXD

Pattern : [GAVLIMPFW]-D-X-X-X-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-[GAVLIMPFW]-X-[GAVLIMPFW]-X-X-F-X-X-Y-X-X-D

## Gleitklammer Region 2:

PCNA\_HUMAN : LKYYLAPKIE  
PCNA\_METJA : LTFLLLAPRIE  
MTH1312 : LSFLLLAPRIE  
PHLA008 : LTFLLLAPRVE  
AF0335 : VEYILAPRIE

Konsensus : \$XXXLAP&\$E

Pattern : [GAVLIMPFW]-X(3)-L-A-P-[KRHDE]-[GAVLIMPFW]-E

Abbildung 5:

## Koppelnde Untereinheit:

PHBN023 : FLEWLNGYVESKEEEEIVSRIRYLIAGDVVD  
 PfuORF2 : FLEWLNGNVENTKEEEEIVSRVKYLIAGDVVD  
 DPD2\_HUMAN: LVDVVTGQLGDEGEQCSAAHVSERVILAGNLLS  
 MTH1405 : FVKWINGDFGSEEQRSLAADVKYLVVAGDIVD  
 AF1790 : FVRWLKGEVGGKKSQNLAEKVKYIVIAGDIVD  
 MJ0702 : FIRFLNGDVDNEEEKVVSRLKYICIAAGDLVD

Konsensus : F\$XX\$XGXXXXXXXXXXXXXX\$XY\$X\$AGD\$\$D  
 L R N S

Pattern : [FL]-[GAVLIMPFW]-X-X-[GAVLIMPFW]-X-G-X(13)-[GAVLIMPFW]-X-[YR]-[GAVLIMPFW]-X-[GAVLIMPFW]-A-G-[DN]-[GAVLIMPFW]-[GAVLIMPFW]-[DS]

Abbildung 6:

## Gleitklammerlader:

AC11\_HUMAN : CNYLSKIIIPALQSRCRTRFRGPL  
 AF2060 : CNYVSRIIEPIQSRCAVFRFKPV  
 MTH0241 : CNYSSKIIDPIQSRCAIFRFLPL  
 PHBN012 : CNYSSKIEPIQSRCAIFRFRPL  
 MJ1422 : CNYPSKIIIPPIQSRCAVFRFSPL

Konsensus : CNYXS&IIX\$\$QSRCXXFRFXP\$

Pattern : C-N-Y-X-S-[KRHDE]-I-I-X-[GAVLIMPFW]-[GAVLIMPFW]-Q-S-R-C-X-X-F-R-F-X-P-[GAVLIMPFW]

Abbildung 7:

## Gleitklammerlader 2:

AC15\_HUMAN : KAALLSGPPGVGKTTTASLV  
MJ0884 : KPILLVGPPGCGKTTLAYAL  
AF1195 : KPLLLAGPPGVGKTSLALAL  
MTH0240 : KPLLLVGPPGTGKTTLAHII  
PHBN013 : KALLLAGPPGSGKTTVYAL

Konsensus : KXXLLXGPPGXGKT\$X\$XX\$

Pattern : K-X-X-L-L-X-G-P-P-G-X-G-K-T- [STNQYC] -X- [GAVLIMPFW] -X-X-  
[GAVLIMPFW]

Abbildung 8:

## Elongationprotein 1

DPOD\_HUMAN : DVITGYNIQNFDPYLISRA  
DPOL\_ARCFU : DIIVGYNQDAFDWPYLRKRA  
MTH1208 : DILVGYNQSDNFDFPYITRRA  
PHBT047 : DVIITYNGDNFDFFPYLLKRA  
MJ0885 : DVIITYNGDNFDFFPYLKARA

Konsensus : D\$XXYNXXXFDXPY\$XXRA

Pattern : D- [GAVLIMPFW] - [GAVLIMPFW] -X-X-Y-N-X-X-F-D-X-P-Y-  
[GAVLIMPFW] -X-X-R-A

Abbildung 9:

## Elongationsprotein 2

AF1722 : AVRTAVAIMTEGVVAAPIEGIARVRI  
 MJ1630 : AVRTALAVLTEGIVAAPLEGIAADVKI  
 PfuORF3 : AVRTALAILTEGIVSAPLEGIAADVKI  
 MTH1536 : ALRTALAILTEGVVAAPLEGIAARVRI  
 PHBN021 : AVRTALAILTEGVVSAPIEGIASVKI

Konsensus: A\$RTA\$A\$\$TEG\$VXAP\$EGIAXV&I

Pattern : A-[GAVLIMPFW]-R-T-A-[GAVLIMPFW]-A-[GAVLIMPFW]-[GAVLIMPFW]-  
 T-E-G-[GAVLIMPFW]-V-X-A-P-[GAVLIMPFW]-E-G-I-A-X-V-[KRHDE]-I

Abbildung 10:

## Elongationsprotein (Eubakteria)

DP3A\_ECOLI : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMPD  
 DP3A\_SALTY : VPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMPD  
 BB0579 : IPVGAGRGSAGSIIVAYALRITDIDPLKYNLLFERFLNPERISMPD  
 DP3A\_HELPY : IPVGPGRGSAAGSLVAFALKITDIDPLKYDLLFERFLNPERISMPD  
 AA50 : IPVGPGRGSAGGSLVAYAIGITDVDPPIKHGFLFERFLNPERVSMPD

Konsensus : \$PVG\$GRGSX\$GS\$VAXA\$XITD\$DP\$XXX\$LFERFLNPER\$SMPD

Pattern : [GAVLIMPFW]-P-V-G-[GAVLIMPFW]-G-R-G-S-X-[GAVLIMPFW]-G-S-[GAVLIMPFW]-V-A-X-A-[GAVLIMPFW]-X-I-T-D-[GAVLIMPFW]-D-P-[GAVLIMPFW]-X-X-[GAVLIMPFW]-L-F-E-R-F-L-N-P-E-R-[GAVLIMPFW]-S-M-P-D

Abbildung 11:

## Gleitklammer (Eubakteria) Region 1:

AAPOL3B : KLVITGGKSTYKLPTAPAEDFP  
DP3B\_ECOLI : RMLVRSGRSRFSLSTLPAADFP  
S.TYPHIM. : RMLVRSGRSRFSLSTLPAADFP  
DP3B\_PROMI : RLLVRSGRSRFSLSTLPASDFP  
DP3B\_PSEPU : KLLVKAGRSRFTLSTLPANDFP  
DP3B\_STRCO : RATVVCGSSRFTLHTLPVEEYP

Konsensus : &XX\$XXGXSXXXLXT\$P\$X&XP

Pattern : [KRHDE]-X-X-[GAVLIMPFW]-X-X-G-X-S-X-X-X-L-X-T-  
[GAVLIMPFW]-P-[GAVLIMPFW]-X-[KRHDE]-X-P

## Gleitklammer (Eubakteria) Region 2:

AAPOL3B : IIMPMRVR  
DP3B\_ECOLI : VVMPMRL  
S.TYPHIM. : VVMPMRL  
DP3B\_PROMI : VVMPMRL  
DP3B\_PSEPU : VVMPMRL  
DP3B\_STRCO : LIMPVRL

Konsensus : \$\$MP\$R\$

Pattern : [GAVLIMPFW]-[GAVLIMPFW]-M-P-[GAVLIMPFW]-R-[GAVLIMPFW]

Abbildung 12:

	20	40	
PCNA_HUMAN	: LINEACWDI SSSGVNLQSMDSHVSLVQLTLRSEGFDTYRCDRNLAMGVNLTSMSKIL : 58		
PCNA_METJA	: LLDEICFEVDEEG IKA SAMDP SHVALVSLEIPRLAFEEYEADS - HDIG IDLEAFKKVM : 57		
MTH1312	: IVDEVQIQLSAEGLRLDALDRSHITYVHLELKAELFDEYVCDEPERINVDTTEELMKVL : 58		
PHLA008	: LIDEAAFKVTEEG ISMRAMDPSRVVLIDLNL PSSIFS KYEV DGEETIGVNMDHLKKVL : 58		
AF0335	: LVSEARIH FLEKGLHSRAVDPANVAMVIVDIPKDSFEVYN IDEEKTIGVDMDRIFDIS : 58		
	60	80	100
PCNA_HUMAN	: KCAGNEDI ITLRAEDNADTLALVFEAPNQE KVSDYEMKLM DLDVEQLG -- IPEQEYSC : 114		
PCNA_METJA	: NRAKAKDRLILEDEEKKNKLNVIFENTG -- KRKFSLALLDISASSVK -- VPEIEYPN : 110		
MTH1312	: KRAKANDR VILSTDEG - N - LIIQFEGEA -- VRTFKIRLIDIEYETPS -- PPEIEYEN : 109		
PHLA008	: KRGKAKDTLILRKGEE - NFLEISLQGTA -- TRTFRLPLIDVEEIEVE -- LPDLPYTA : 110		
AF0335	: KSISTKDLV LELIVEDE - STLKV KFGSVE ---- YKVALIDPSAIRKEPRIPEELPA : 109		
	120	140	160
PCNA_HUMAN	: VVKMPSGEFARICRDL SHIGDAVVI SCAKDGKFSASGELGNGN IKL SQT SNVDKEEE : 172		
PCNA_METJA	: VIMIKGDAFKEALKDADLFS DYVILKADEDKFV I HAKGDLNENE AIFEKDSSA ----- : 163		
MTH1312	: EFEVPPFQLLKDSIADIDIFS DSDK I FTRV DEDRF I ASAEGEFGDAQIEY ----- : 156		
PHLA008	: KVVVLGEVILKEAVKDA SLVSDS I KFMAKENE FIMRAEGETQEV EVKLTLEDEG ----- : 163		
AF0335	: KIVMDAGEFKKA IAAADKIS DQV I FRS DKEGFRIEAKGDVDSIVFHMT - ET ----- : 159		
	180	200	220
PCNA_HUMAN	: AVTIEMNEPVQLTFAL RYLNFFT KATPLSSTV LMSADVPLVVEYK IADM - GHLKYY : 229		
PCNA_METJA	: IISLEVKEAKSAFNLDYLM DMVKGVSSGDIKIYLGNDMPLKLEY SIAG -- VNLTFL : 219		
MTH1312	: LHGERIDKPARSIYSLDKIKEMLKADKFSETAIINLGDDMPLKLT LKMASKEGELSFL : 214		
PHLA008	: LLDIEVQEETKSAYGVSYLADMVKG ICKADEVTMRFGNEMPMQMEYYIRDE - GRLTFL : 220		
AF0335	: ELIEFNGGEARSMF SVDYLKEFCKVAGSGD LLIHGT NYPVRLV FELVGGRAKVEYI : 217		
PCNA_HUMAN	: LAPKIED : 236		
PCNA_METJA	: LAPRIEG : 226		
MTH1312	: LAPRIEA : 221		
PHLA008	: LAPRVEE : 227		
AF0335	: LAPRIES : 224		

Abbildung 13:

	20	40	
AAPOL3B	MRVKVDREEEVLKKARESTEKKAAPILANFLLSAKEENLIVRATDLENVLVSVK	:	58
DP3B_ECOLI	MKFTVEREHLKPLQQVSGPLGGRPTLPILGNLLQVADGTLSTLTGTDLEMEMVARVA	:	58
S.TYPHIM.	MKFTVEREHLKPLQQVSGPLGGRPTLPILGNLLQVADGTLSTLTGTDLEMEMVARVT	:	58
DP3B_PROMI	MKFIIIEREQLKPLQQVSGPLGGRPTLPILGNLLKVTENTLSTLTGTDLEMEMMARVS	:	58
DP3B_PSEPU	MHFTIQREALLKPLQLVAGVVERRQTLPVLSNVLVQGQQQLSTLTGTDLEVELVGRVQ	:	58
DP3B_STRCO	MKIRVERDVLAEAVAWAARSLPARPPAPVLAGLLLKAEEGQLSLSFDYEV SARV SVE	:	58
	60	80	100
AAPOL3B	GEVEEE-GEVCVHSQKL-YD1VKNL-NSAYVYLHTEGEKLVITGGKSTYKLPTAPAEDF	:	114
DP3B_ECOLI	LVQPHEPGATTVPARKFFDICRGLPEGAEIAVQLEGERMLVRSGRSRFSLSTLPAADF	:	116
S.TYPHIM.	LSQPHEPGATTVPARKFFDICRGLPEGAEIAVQLEGDRMLVRSGRSRFSLSTLPAADF	:	116
DP3B_PROMI	LSQSHEIGATVPARKFFDIWRGLPEGAEISVELGDRLLVRSGRSRFSLSTLPASDF	:	116
DP3B_PSEPU	LEEPAAEPGEITVPARKLMDICKSLPNDALIDIKVDEQKLKVAGRGRFTLSTLPANDF	:	116
DP3B_STRCO	AEIEEE-GTFLVSGRLADISRAL-PNRPVEISTDGVRAVVCGSSRTLHTLPVEEY	:	114
	120	140	160
AAPOL3B	PEFPEIVE-GGETLSGNLLVNGIEKVEYAIAKEEANIALQGMYLRYEDRIHFVGSDG	:	171
DP3B_ECOLI	PNLDDWQSEVEFTLPQATMKRLIEATQFSMAHQDVRYYLNGMLFETEGEELRTVATDG	:	174
S.TYPHIM.	PNLDDWQSEVEFTLPQATMKRLIESTQFSMAHQDVRYYLNGMLFETEGSELRTVATDG	:	174
DP3B_PROMI	PNLDDWQSEVEFTLPQATLKRLIESTQFSMAHQDVRYYLNGMLFETEMTELRTVATDG	:	174
DP3B_PSEPU	PTVEEGPGSLTCNLEQSKLRRRLIERTSFAMAQQDVRYYLNGMLLEVSRTLRAVSTDG	:	174
DP3B_STRCO	PALPQMPE-ATGTVPGEVFASAVQVAAAGRDDTLPVLTGVRIEIEGDSVTLASTDR	:	171
	180	200	220
AAPOL3B	HRLALYEPLEGEFSKEL----LIPRKS LKV LKLITGIEDVNIEKSE---DESFAYFS	:	221
DP3B_ECOLI	HRLAVCSMPIQQLSPLS--HSVIVPRKGVIELMRMLDG-GDNPLRVQI---GSNNIRAH	:	226
S.TYPHIM.	HRLAVCSMPLEASLPS--HSVIVPRKGVIELMRMLDG-GENPLRVQI---GSNNIRAH	:	226
DP3B_PROMI	HRLAVCAMDIQQLSPLG--HSVIVPRKGVIELMRMLDGSGESLLQLQI---GSNNLRAH	:	227
DP3B_PSEPU	HRLALCSMSAPIEQEDR-HQIVIVPRKGILELARLTD-PEGMVSIVL---GQHHIRAT	:	227
DP3B_STRCO	YRFAVREFLWKPNPDI S AVALVPAKTLQDTAKALTSGDQVILALSGSGAGEGLIGFE	:	229
	240	260	280
AAPOL3B	TPEWKLAVRLLEGEPDYMNSVIPPEEFSAEVLFETEEVLKVLKRLKALSEGKVFPVKIT	:	279
DP3B_ECOLI	VGDFIFTSKLVDGRFPDYRRVLPKPNPDKHLAEGCDLLKQAFARAIALSNEKFRGVRLY	:	284
S.TYPHIM.	VGDFIFTSKLVDGRFPDYRRVLPKPNPDKHLAEGCDILKQAFARAIALSNEKFRGVRLY	:	284
DP3B_PROMI	VGDFIFTSKLVDGRFPDYRRVLPKPNPTKTVIAGCDILKQAFSRAAILSNEKFRGVRLIN	:	285
DP3B_PSEPU	TGEFTFTSKLVDGKFPDYERVLPKGGDKLVVGRDQALREAFSRTAILSNEKYRGIRLQ	:	285
DP3B_STRCO	GAGRRTTTRLLEGDLPKYKTLFPTEFNSAVIETAPFVEAVKRALVA-ERNTPVRLS	:	286
	300	320	340
AAPOL3B	LSENLAIFEFADPEFGEAREEIEVEYTGEPIFNGKYLMEALDAYDSERVWFKFTT	:	337
DP3B_ECOLI	VSENQLKITANNPEQEAAEIELDVTYSGAEMEIGFNVSYVLDVLNALKCENVRMLT-	:	341
S.TYPHIM.	VSENQLKITANNPEQEAAEIELDVSYGGTEMEIGFNVSYVLDVLNALKCETVRMLT-	:	341
DP3B_PROMI	LTNGQLKITANNPEQEAAEIELDVSYGGTEMEIGFNVSYLLDVLNTLKCEEVKLLT-	:	342
DP3B_PSEPU	LAAGQLKIQANNPEQEAAEELISVDYEGSSLEIGFNVSYLLDVLGVMTTEQVRLILS-	:	342
DP3B_STRCO	FEQGVILIEAGSSDDAQAVERVDAQLEGDDISIAFNPTFLDGLSAIDSPVAQLSFTT	:	344
	360	380	
AAPOL3B	---PDTATLLEAEDYEK-EPYKCIIIMPMRV--	:	363
DP3B_ECOLI	---DSVSSVQIEDAAS-QSAAYVVMRML--	:	366
S.TYPHIM.	---DSVSSVQIEDAAS-QSAAYVVMRML--	:	366
DP3B_PROMI	---DAVSSVQENVAS-AAAAYVVMRML--	:	367
DP3B_PSEPU	---DSNSSALLQEAGN-DDSSYVVMRML--	:	367
DP3B_STRCO	STKPALLSGRPAPDAEADAEAYKYLIMPVRLSG	:	376

## Abbildung 14:

	* 20	* 40	* 60	
AC11_HUMAN	: ICNYLSKIIPALQSRCTRFRGPLTPELMVPRLEHVVEEKVDISEDMKALVTLSSGDM			: 60
AF2060	: SCNYVSRIIEPIQSRCAVFRFKPVPKAEAMKKRLLEICEKEGVKITEDGLEALIYISGGDF			: 60
MTH0241	: SCNYSSKIIDPIQSRCAIFRFLPLKGHQIICKLEYIAEKENLEYAHALETIVYFAEGDL			: 60
PHBN012	: SCNYSSKIIIEPIQSRCAIFRFRPLRDEIAKRLRYIAENEGLELTEEGLQAILYIAEGDM			: 60
MJ1422	: NCNYP SKIIPPIQSRCAVFRFSPLKKEDIAKKLKEIAEKG LNLTESGLEAIYVSEGDM			: 60

	* 80	* 100	* 120	
AC11_HUMAN	: RRALNIQSTNMAFGKVTEETVYTCGHPLKSDIANILDWMLNQDFTTAYRNITELKTLK			: 120
AF2060	: RKA INALQAAAIGEVVADT IYQITATARPEEMTEL IQTALKGNFMEARELLDRIMVEY			: 120
MTH0241	: RKA INLLQSAASLGEKITESSIIYDVVSARPKDVRKMIKTILDGKFMEARDMLREIMVLQ			: 120
PHBN012	: RRA INILQAAAALDKKIDENVF MVASRARPEDIREMMLLALKGNFLKAREKLREILLKQ			: 120
MJ1422	: RKA INVLQTAALSDVIDDEIVYKVSSRARPEEVKKM MELALDGKFMEARDLLYKLMVEW			: 120

	*	
AC11_HUMAN	: GLALHDILTEIHLFVHRVD	: 139
AF2060	: GMSGEDIVAQLFREIISMP	: 139
MTH0241	: GISGEDMVTQIYQELSRLA	: 139
PHBN012	: GLSGEDVLIQMHKEVFNL P	: 139
MJ1422	: GMSGEDILNQMFREINSLD	: 139

## Abbildung 15:

	* 20 *	40 *								
AC15_HUMAN	: LLWVDKYKPTSLKTIIGQQGDQSCANLRLRNWQKSSSEDKHAAKFGKFSGKDDG	:	58							
MJ0884	: LSWVEKYRPKSLKDVG--HEKVEKEKLKTVIESYLNGET-----	:	37							
AF1195	: MLWVEKYRPKTLVEVVA---DKSITRVIKWAWSKRG-----	:	35							
MTH0240	: MSWTEKYRPGSFDEVVG---NQKVIAEIKEWIKAWKAGKP-----	:	37							
PHBN013	: VPWIEKYPRKLSEIVN---QEQALESKRAWIESWLHGNPPK-----	:	39							
	60 *	80 *	100 *							
AC15_HUMAN	: SSFKAALLSGPVGKTTTASLVCQELGYSYVELN ASDTRSKSSLKAI	VAESLNNTSI	:	116						
MJ0884	: -PKPILLVPGPPGCGKTTLAYALANDYGF	EVIELNA	SDKRNSSAIKKVVGHAATSSSI	:	93					
AF1195	: -SKPLLALAGPPVGKTS	LALALANTMGWEA	VELN ASDQRSWRVIERIVGEGAFNETI	:	91					
MTH0240	: -QKPLLVLGVPPGTGKTTLAHIIGKEFS-DTLELNASDRRSQDALMRSAGEASATRSL	:	92							
PHBN013	: -KKALLLAAGPPSGKTTTVYALAH	EYNFEVIELN ASDERTYNIARYV-QAAYTMDI	:	94						
	120 *	140 *	160 *							
AC15_HUMAN	: KGFYSNGAASSVSTKH	ALIMDEV	DGMAGNEDRGGIQELIGLIK-HTKIP	IIICMCNDRN	:	173				
MJ0884	: FG-----KKF-LIVLDEV	DVG	ISGKEDAGGVSELIKVIK-KAKN	PIIL	TANDAY	:	139			
AF1195	: SDEGEF-LSSRIGKLKLI	ILDEV	DNIHKKEDVGGEALIRLIK	RKPAQPLI	LIANDPY	:	148			
MTH0240	: FN-----HDLKLI	ILDEV	DVG	IHGNE	DRGGVQAINRIIK-ESRHPMVLT	TANDPY	:	139		
PHBN013	: MG-----KRRKII	FLDEADNIE--P--SGAPEIAKLID-KARNPI	IMAANHYW	:	137					
	180 *	200 *	220 *							
AC15_HUMAN	: HPKIRSLVHYCFDLRFQPRVEQIKGAMMSIAFKEGLKIPPPAMNEII	LGANQDIRQV	:	231						
MJ0884	: APSIRSLLPYVEVIQLNPVHTNSVYKVLK	KIAEKEGLDVD	DKTLK	MIQHSAGDLRSA	:	197				
AF1195	: KLSP-ELRNLC	EMINFKRLTKQQVARV	LERIALKEGI	KVDKSVLLKIAE	ENAGGDLRAA	:	205			
MTH0240	: SKRLQSIKPRCRVLNLRKVHTSSIAALRRICRAEGIECPDDVL	RELAKRSRGDL	DLRAA	:	197					
PHBN013	: EVPK-EIRDRAELVEYKRLNQRDV	ISALVRILKREGIT	VPKEILTEIAKRSSGDLRAA	:	194					
	240 *	260 *	280 *							
AC15_HUMAN	: LHNLSMWCARSKALTYDQAKADSHRAKKDI	MGPF	DVARKVFAAGEETAHMSLVDKSD	:	289					
MJ0884	: INDLEALALSGDLSYEAAQKLP-----DRKREANIF	DALR	VILKTTHYGIATTALMN--	:	249					
AF1195	: INDFOALAELEGKEELKPEDVF	L-----KRTQE	KDIFRVMQMF	IKTKNP	AVYNEAML---	:	256			
MTH0240	: INDLEAMAEGERIGEELLKMG-----EKDATSNLF	DAVRAVLK	SRDVSKVREAMR	--	:	248				
PHBN013	: INDLQTIVAGG--YEDAKYVLA-----YRDVEKTVF	QSLGMVF	SSDNAKRAKL	ALMN--	:	244				
	300 *	320 *	340 *							
AC15_HUMAN	: LFFHDYSIAPLFVQENYI	HVKPV	AAGGDMKKHMLLSRAADSICDGLVDSQIRSKQ	N	:	347				
MJ0884	: -----VDETPDV	VIEWIAENVP	KEYEKPEEVARAF	EYELSKADRYLGRVMRRQ	N--YSF	:	300			
AF1195	: -----LDESPEDV	IHWV	DENLP	LEYSG-VELVNAYE	ALSRADIFLGRVRRQ	F--YRL	:	306		
MTH0240	: -----VDDDP	TLV	LEFIAENVP	REYEKPN	EISRAYDMLSRADIFFGRAVR	TRN--YTY	:	299		
PHBN013	: -----LDMSPDEF	LWV	DENIPH	MYLKPEEMARAYEA	ISRADIYLGRAQRT	GN--YSL	:	295		
	360 *	380 *	400 *							
AC15_HUMAN	: WSLLPAQAIYASVLP	GELMRGYMTQFPTF	PSWL	GKHSSTGKHDRIVQDLALHMSLRTY	:	405				
MJ0884	: WKYATT	LMTAGVALSK	DEK	YRKWTPSY-PKIFRLLTK	KAEREILNKILKKIGEK	TH	:	357		
AF1195	: WKYASYLMT	VGQQMKEEP	KKG	FTRYRR-PAVWQMLFQLRQK	REMTRKILEKIG	KYSH	:	363		
MTH0240	: WRYASEL	MGPGVALAKD	KTYRK	FVRYTG-SSSFRILG	KTRKQ	SLRDSVAAKMAG	KMH	:	356	
PHBN013	: WKYAI	DMMTAGV	VAVAG	TK-KKGFAK	FYP-PNTL	KMLAESKEERS	IRD	SIKKIMKEMH	:	351
	420 *									
AC15_HUMAN	: SSKRTVNMDYLSLLR	:	420							
MJ0884	: TSSKRAR-FDLQMLK	:	371							
AF1195	: LSMRKARTEMFPVIK	:	378							
MTH0240	: ISPKV	AI-SMFPYME	:	370						
PHBN013	: MSKLEAL-ETMKILR	:	365							

## Abbildung 16:

	* 20 *	* 40 *	
PHBN023	: IFEVEQDTDRVKVFLPKDSED-YREALKVLPDAVAFKGKVYSKRG-IFFANRFYLPDV	: 56	
PfuORF2	: IFEIEDLTGKVKVFLPKDSED-YREAFKVLPAVVAFKGVYSKRG-ILYANKFYLDPDV	: 56	
DPD2_HUMAN	: -LVLEDELQRIKLIKGTIDVSK-----LVTGTVLAVFGSVRDDGKFLVEDYCFADLA	: 50	
MTH1405	: IIELEDDTGEISVVVHNENHKLFEKSEKIVRDEVVGVHGTTKGR--FVVASEIFHPGV	: 56	
AF1790	: YIRLEDTTGTTICVATGKNAE---VARELLGDEVIGVTGLLKGS--SLYANRIVFPDV	: 53	
MJ0702	: IVRIEDTEDEATLILPKEKIEAGKIPDDILLDEVIGAIGTVSKSGSSIYVDEIIRPAL	: 58	
	60 * 80 *	100 *	
PHBN023	: PLYRK-QKPPLEEKVYAVLTSIHAVGSK--EFCEKAFIKFLEWLNGYVESKEEEEIVS	: 111	
PfuORF2	: PLYRR-QKPPLEEKVYAILISDIHVGSK--EFCENAFIKFLEWLNGNVTKEEEEIVS	: 111	
DPD2_HUMAN	: PQKP--APP-LDTDRFVLLVSGLGLGGGGESLLGT-QLLVDVVTGQLGDEGEQCSAA	: 104	
MTH1405	: PRIQ--EK--EMDFSVAFISDVHIGSQ--TFLEDAFMFKVKWINGDFGSEEQRLAA	: 107	
AF1790	: PINGNGEK--KRDFYIVFLSDTHFGSK--EFLEKEWEMFVRWLKGEVGGKKSQNLAE	: 106	
MJ0702	: PPKEP-KR--IDEEIYMAFLSDIHVGSK--EFLHKEFEKFIRFLNGDVDNEEKKVVS	: 111	
	120 * 140 *	160 *	
PHBN023	: RIRYLLIAGDVVD-GIGIYP-GQYSDLIIPDIFDQYEALANLLSNVP---KHITIFI	: 163	
PfuORF2	: RVKYLILIAGDVVD-GVGVYP-GQYADLTIPDIFDQYEALANLLSHVP---KHITMFI	: 163	
DPD2_HUMAN	: HVSRVILAGNLLSHSTSQRDSINKAKYLTKKTQAASVEAVKMLDEILLQLSASPVVDV	: 162	
MTH1405	: DVKYLVVAGDIVD-GIGIYP-GQEKELLIRDIHEQYEEAARLFGDIR---SDIKIVM	: 159	
AF1790	: KVKYIVIAGDIVD-GIGVYP-GQEDDLAISDIYQGYEFAASHLDEIP---KEKIIIV	: 158	
MJ0702	: RLKYICIAIGDLVD-GVGVYP-GQEEDLYEVDDIEQYREIAMYLDQIP---EHISIII	: 163	
	180 * 200 *	220 *	
PHBN023	: GPGNHDAARPAIPOPEFYEEYAKPLVYKLKNTVIISNPAPIVRLHGRDFLIAHGRGIEDV	: 221	
PfuORF2	: APGNHDAARQAIPQPEFYKEYAKPIYKLKNAVIISNPAPIVRLHGRDFLIAHGRGIEDV	: 221	
DPD2_HUMAN	: MPGEFDPTNYTLPQQPLHPCMFPPLATAYSTLQLVTNPYQATIDGVRFLGTSGQNVSDI	: 220	
MTH1405	: IPGNHDSSRIAEPQPAIPEEYAKSLYSIRNIEFLSNPSLIVSLDGVRTLIIYHGRSFDDM	: 217	
AF1790	: SPGNHDAVRQAEPQPAFEGEI-RSLFP-KNVEHVGNPAYVDIEGVKVLIIYHGRSIDDI	: 214	
MJ0702	: SPGNHDAVRPAEPQPKLPEKITKLFNR-DNIYFVGNPCTLNIHGFDLILYHGRSFDDL	: 220	
	240 * 260 *	280 *	
PHBN023	: VSFVPGLTHHKPGLPMVELLKMRHLAPTFGGKVPIAPDPE-DLLVIEVPDLVQMGHV	: 278	
PfuORF2	: VGSVPGLTHHKPGLPMVELLKMRHVAPMF GGKVPIAPDPE-DLLVIEVPDVVHMGHV	: 278	
DPD2_HUMAN	: FRYS---SMEDHLEILEWTLRVRHISPTAPDTLGCYPFYKTDFIFPECPhVYFCGNT	: 275	
MTH1405	: AMSVNGLSHERSDSLIMEELLEKRRHLAPIYGERTPLASEIE-DHLVIEDVPVLHTGHV	: 274	
AF1790	: ISKIPRLSYDEPQKVMEEELLKRRHLSPYGGRTPLAPERE-DYLVIDVPDILHCGHI	: 271	
MJ0702	: VQQIRAAASYENPVTIMKELIKRRLLCPTYGGRCPIAPEHK-DYLVIDRDIDILHTGHI	: 277	
	300 * 320 *		
PHBN023	: HVYDTAVYRG----VQLVNSATWQAQTEFQKMVNIVPTPGLVPIVD	: 320	
PfuORF2	: HVYDAVYRG----VQLVNSATWQAQTEFQKMVNIVPTPAKVPVD	: 320	
DPD2_HUMAN	: PSFGSKIIRGPEDQTLLVTVPDF-SATQTACLVNLR-SLACQPISF	: 320	
MTH1405	: HINAYKKYKG----VHLINSGTFSQTEFQKCIYNIVPTCGQVPLN	: 316	
AF1790	: HTYGTGFYRG----VFMVNSSTWQAQTEFQKKVNLNPMPGNVAVYR	: 313	
MJ0702	: HINGYGIYRG----VVMVNSGTFQEQTDFQKRMGJISPTPAIVPIIN	: 319	

## Abbildung 17:

	* 20 *	* 40 *		
DPOD_HUMAN	: KVQSYEKEEDLLQAWSTFIRIMDPDVITGYNIQNFDLPLYLISRAQTLKVQTFPFLGRV	: 58		
DPOL_ARCFU	: EJILTGDERKIISDFVKLIKSYDPPDIIVGYNQDAFDWPYLRKRAERWNIPLD--VG--	: 54		
MTH1208	: FVEVVEDERELLERFAEIVIDKKPDILVGYNSDNFDFPYITRRAAILGAELD--LG--	: 54		
PHBT047	: YVEVVSSEREMIKLIRVIKEKDPDVIIITYNGDNFDFFPYLLKRAEKGIGKLL--LG--	: 54		
MJ0885	: NIEVVKNEKELIKKIIETLKEY--DVIYTYNGDNFDFFPYLKARAKIYGIDIN--LG--	: 52		
	* 60 *	* 80 *	* 100 *	
DPOD_HUMAN	: AGLCSNIRDSSFOSKQTGRRTDKVVSMMGRVQMDMLQVLLREYKRSHTLNAVSFHFL	: 116		
DPOL_ARCFU	: -----RDGSNVVPRGG-----RKPITGRLNVDLYDIAMRISDIKKLENVAEFLG	: 100		
MTH1208	: -----WDGSKIRTMRRG-FANATAIKGTVHVDLYPVMMRYYMNLDRYTLERVYQELF	: 104		
PHBT047	: -----RDNSEPKMQKMG-DSLAVEIKGRIHFDLFPVIRRTINLPTYTLLEAVYEAF	: 104		
MJ0885	: -----KGGEELKIKRGG-MEYRSYIPGRVHIDLPISSRLKLTKTLEDVVYNLF	: 102		
	* 120 *	* 140 *	* 160 *	
DPOD_HUMAN	: GEQKE-DVQHSIIITDLQNGNDQTRRRLAVYCLKDAYLPLRLLERLMVLVNAVEMARVT	: 173		
DPOL_ARCFU	: TKIEIADIEAKDIYRYWSRGE--KEKVLNYARQDAINTYLIAK--ELPPMHYELSKMI	: 154		
MTH1208	: GEEKI-DLPGDRLWEYWDRDEL-RDELFYSLDDVVATHRIAEG-KILPLNLELTRLV	: 158		
PHBT047	: GKPKE-KVYADEIAKAWETGEG-LERVAKYSMEDAKVTYELGR--EFFPMEAQLARLV	: 158		
MJ0885	: GIEKL-KIPHTKIVDYWANND--KTLIEYSLQDAKYTYKIGK--YFFPLEVMFSRIV	: 154		
	* 180 *	* 200 *	* 220 *	
DPOD_HUMAN	: GPVLSYLLSRGQQVKVVSQQLRQAMHEGLLMPVVKSE-----GGEDYTGATVIEPLK	: 225		
DPOL_ARCFU	: RLPVDDVTRMCRGKQVDWLLSEAKKIGEIAPNPPE-----HAESYEGAFVLEPER	: 205		
MTH1208	: GQPLFDISRMATGQQAEWFLVRKAYQYGELVPNKPSQSDFSSRRGRRAVGGYVKEPEK	: 216		
PHBT047	: GQPVWDVSRSSSTGNLVEWFLLRAYERNELAPNKPDEKEYERRLRESYEGGYVKEPEK	: 216		
MJ0885	: NQTPFEITRMSSGQMVEYLLMRAFKENMIVPNKPDEEEYRRRLTTYEGGYVKEPEK	: 212		
	240			
DPOD_HUMAN	: GYYDVPIATLDFS	: 238		
DPOL_ARCFU	: GLHEN-VACLDFA	: 217		
MTH1208	: GLHEN-IVQFDFR	: 228		
PHBT047	: GLWEG-IVSLDFR	: 228		
MJ0885	: GMFED-IIISMDFR	: 224		

## Abbildung 18:

	20	40	60	
AF1722	DTIKGVKGMTSKTKI PERLEKG ILRVKHGVFVKDG TARP DATDLP ITHFKP AEIGVSVEKLR			: 63
MJ1630	GDVKCIKGMTSKQK IVEPLEKAI LR AINEVYVFKDG TTRFDCTDVPVTHFKPNEINVTEKLR			: 63
PfuORF3	DKLGKVGMMTSGWKIAE PLEKGLLRAKNEVYVFKDG TIRFDATDAPITHFRP REIGVSVEKLR			: 63
MTH1536	DEIKGV EGMIS A EKFPE PLEKGLI RAKNDVYTFKDATIRHDSTDLPLTHFTP REIGVSVERLR			: 63
PHBN021	DKLGKVGMMTSGWKI MPEPLEKGLLRAKNDVYVFKDG TIRFDATDAPITHFRP REIGVSVEKLR			: 63

	80	100	120	
AF1722	ELGYERDYKGAE LKNEQ IVELKPQDVILPKSGA EYLLRVANFIDLLVFKFYKMEPFYNAKSV			: 126
MJ1630	ELGYDKDIYGNELVDGEQVVELKPQDVII PESCAEYFVKVANFIDDLLEKFYKVERFYNVKKK			: 126
PfuORF3	ELGYTHDFEGKELVSEDQ IVELKPQDVILSKEAGKYL RVARFV DLLLEKFYGLPRFYNAEKM			: 126
MTH1536	ELGYTRDCYGD ELED EDEQ ILELRVQDV VISEDCADYLVRVANFV DLLLERFYDLERFYNVKTR			: 126
PHBN021	ELGYTHDFEGNPLVSEDQ IVELKPQDIILSKEAGKYL KVAKFV DLLLEKFYGLPRFYNAEKM			: 126

	140	160		
AF1722	EDLIGHLVIGLAPHTSAGV LGRIIGFS DVLAGYAH PYFHA KRR			: 170
MJ1630	EDLIGHLVIGMAPHTSAGMVGRIIGYTKANVGYAH PYFHA KRR			: 170
PfuORF3	EDLIGHLVIGLAPHTSAGIVGRIIGFVDALVGYAH PYFHA KRR			: 170
MTH1536	EDLVGH LIAGLAPHTSAAVLGRIIGFTGASAC YAH PYFHS A KRR			: 170
PHBN021	EDLIGHLVIGLAPHTSAGIVGRIIGFVDALVGYAH PYFHA KRR			: 170

## Abbildung 19:

	20	40	60	
DP3A_ECOLI	ELQVINQMGFPGYFLIVMEFIQWSKDNGVPVG PGRGSGAGSLVAYALKITDLDPLEFDLL			: 60
DP3A_SALTY	ELQVINQMGFPGYFLIVMEFIQWSKDNGVPVG PGRGSGAGSLVAYALKITDLDPLEFDLL			: 60
BB0579	ELSVIIGMGFEGYFLIVWDFIKFAHDNDIPVGAGRGSGAGSIVAYALRITDIDPLKYNLL			: 60
DP3A_HELPY	EIEVITNMKFPGYMLIVWDFIRYAKEMGIPVGPGRGSAGSLVAFALKITDIDPLKYDLL			: 60
AA50	ELEVINKMGFAGYFLIVQDFINWAKKNDIPVGPGRGSAGGSVAYAIGITDVPDIKGFL			: 60

	80	100	120	
DP3A_ECOLI	FERFLNPERVSMPDFDVDFCMEKR DQVIEHVADMYGRDAVSQIITFGTMAAKAVIRDVGR			: 120
DP3A_SALTY	FERFLNPERVSMPDFDVDFCMEKR DQVIEHVADMYGRDAVSQIITFGTMAAKAVIRDVGR			: 120
BB0579	FERFLNPERISMPDFDIDFCFEGR DEI I KYVTKYGEDKVAQIITFGTLKPKAVVKDVAR			: 120
DP3A_HELPY	FERFLNPERISMPDIDTDFCQRRRKEII EYMI EKYGKYNVAQVITFNKMLAKGVIRDVAR			: 120
AA50	FERFLNPERVSMPDIDVDFCQDNREK VIEYVRNKYGH DNVQAQIITYNV MKA KQTLRDVAR			: 120

DP3A_ECOLI	VLGHPYGFVDRISKLI PP			: 138
DP3A_SALTY	VLGHPYGFVDRISKLV PP			: 138
BB0579	VLDIPFAESNELTKFI PD			: 138
DP3A_HELPY	VLDMPYKEADDFAKLI PN			: 138
AA50	AMGLPYSTADKLAKLI PQ			: 138

# A thermostable in vitro polymerase complex for template-dependent elongation of nucleic acids in amplification or reverse transcription methods

Patent Number: DE19840771

Publication date: 2000-02-10

Inventor(s): KOBER INGO (DE); MOECKEL GERD (DE); VOSS HARTMUT (DE); KILGER CHRISTIAN (DE)

Applicant(s): LION BIOSCIENCE AG (DE)

Requested Patent:  DE19840771

Application Number: DE19981040771 19980907

Priority Number(s): DE19981040771 19980907; DE19981035653 19980806

IPC Classification: C12N9/12; C07K19/00; C12P19/34; C12N15/62; C12N15/63; C12Q1/68

EC Classification: C07K14/195, C12N9/12B7B7, C12N9/12B7B49

Equivalents:

## Abstract

A thermostable in vitro complex for template-dependent elongation of nucleic acids comprises a thermostable sliding clamp protein, which is connected with an elongation protein that shows thermostable polymerase activity is new. Independent claims are also included for the following: (1) a thermostable accessory in vitro complex characterized in that it contains a sliding clamp protein and a coupling protein, as defined above; (2) a recombinant DNA sequence, which encodes a thermostable in vitro complex or a thermostable accessory complex; (3) a vector that contains recombinant DNA encoding a sliding clamp protein and a coupling protein and/or an elongation protein; (4) a host cell transformed with one or more vectors of (3); (5) a method for the production of a thermostable in vitro complex or thermostable in vitro accessory complex as above; (6) a method for template-dependent elongation of nucleic acids; (7) a method of marking nucleic acids through generating a single break in the phosphodiester bond of the nucleic acid chain and substituting a nucleotide at the break site with a marked nucleotide with the help of a polymerase, where the polymerase is a thermostable in vitro complex as above; and (8) a reagent kit for elongation and/or amplification and/or reverse-transcription and/or sequencing and/or marking of nucleic acids.

Data supplied from the esp@cenet database - I2